NASA TECHNICAL TRANSLATION

NASA TT F-244

c. 1

NASA TT F-244

PROBLEMS OF SOLAR ACTIVITY

by B. M. Rubashev

Main Astronomical Observatory

"Nauka" Publishing House, Moscow-Leningrad, 1964

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • DECEMBER 1964

PROBLEMS OF SOLAR ACTIVITY

By B. M. Rubashev

Translation of "Problemy solnechnoy aktivnosti"

Main Astronomical Observatory
"Nauka" Publishing House, Moscow-Leningrad, 1964

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

For sale by the Office of Technical Services, Department of Commerce,
Washington, D.C. 20230 -- Price $7.00

CONTENTS

FOREWORD

During the past two decades, a number of monographs have been published which discuss problems involved in solar activity and its manifestations. The following are among the monographs which should be mentioned: the second edition of Waldmeier's book "Results and Problems in Solar Research",[1] his earlier book "Sun and Earth",[2] Ellison's book "The Sun and its Influence on the Earth",[3] Gleissberg's book "Sunspot Frequency",[4] Clayton's two-volume monograph "Solar Relations to Weather and Life",[5] Berg's book "Solar-Terrestrial Relationships in Meteorology and Biology",[6] and studies of Soviet authors: "Solar Activity and its Terrestrial Manifestations"[7] and "Indications of Physicogeographic Manifestations of Solar Activity".[8] In addition, in a number of monographs of similar content a certain amount of attention also has been given to the problems involved in solar activity and its geophysical manifestations. Particular mention should be made of Baur's book "Introduction to Macrometeorology".[9]

---

[1]M. Waldmeier. Ergebnisse und Probleme der Sonnenforschung. Akademie Verlag. Leipzig, 1955.
[2]M. Waldmeier. Sonne und Erde. Gutenberg, Zurich, 1946.
[3]M. A. Ellison. Solntse i yego vliyaniye na Zemlyu. Fizmatgiz, Moscow, 1959.
[4]W. Gleissberg. Die Haufigkeit der Sonnenflecken. Akademie Verlag, Berlin, 1952.
[5]H. H. Clayton. Solar Relations to Weather and Life, I, II, Canton, Massachusetts, 1943
[6]H. Berg. Solar-terrestrische Beziehungen in Meteorologie und Biologie. Akademie Verlag, Leipzig, 1957.
[7]M. S. Eygenson, M. N. Gnevyshev, A. I. Ol' and B. M. Rubashev. Solnechnaya aktivnost' i yeye zemnyye proyavleniya. Gostekhizdat, Moscow-Leningrad, 1948.
[8]M. S. Eygenson. Ocherki fiziko-geograficheskikh proyavleniy solnechnoy aktivnosti. Izd. L'vovsk. univ., 1957.
[9]F. Baur. Einfuhrung in die Grosswetterkunde. Dietrich, Wiesbaden, 1948

A characteristic feature of almost all these books is a nonuniform distribution of material by aspects and uneven detail in discussion of individual problems. Many books are devoted entirely to a single aspect of the problem. For example, in Waldmeier's book, "Results and Problems in Solar Research", very little attention is given to the sun-earth problem and the sketchiness of exposition of the principal aspects of this problem in comparison with the detailed discussion of the direct problems of solar activity is immediately noticeable. In the Ellison book, very little attention has been give to the solar cycles, and the author mentions this in the foreword; in addition, the sun-troposphere problem is not considered. On the other hand, the books by H. Berg and Clayton are concerned almost exclusively with this problem. Thus, with rare exceptions, the entire major problem of solar activity and its manifestations has not been considered together in any book, nor in its many aspects in such a way that the principal problems are given equal attention.

Another shortcoming of these works is of an objective character; there is an exceedingly large amount of material based on empirical-statistical investigations, and very little based on the results of physicomathematical work. We term this shortcoming "objective" because this fact reflects the actual status of the matter. There have been incomparably more statistical investigations of every possible type devoted to these problems than there have been theoretical studies. The number of facts awaiting a physicomathematical interpretation is now considerable; although this does not mean, of course, that there is no need for newly observed facts; on the contrary, the need for new facts is now extremely great. However, it must be said that there now is a possibility of reducing to a certain system, all of the numerous results of investigations of the solar cycle, solar active radiation, the mechanisms of the relationship between solar activity and phenomena in the various envelopes of bodies in the solar system, etc.

Such investigations exist (although in limited number), but they are scattered throughout various periodicals, many of which are difficult to obtain. At the same time, the inclusion of these results in a monograph, even in preliminary form, in our opinion would be highly useful. In particular, it would be possible in this manner to find a rational group of physical mechanisms responsible for the manifestation of solar activity in geophysical and other processes. This would facilitate the clarification of the role of solar activity as the singular director of an entire series of processes in the solar system.

Finally, we note another fact, possibly the most important, which has not been adequately treated in the abovementioned monographs. At the present time, an increasing number of geophysicists are recognizing the importance of solar activity and the need for taking it into account when making various kinds of geophysical forecasts. It is

well-known that solar activity must be taken into account when forecasting the state of the geomagnetic field and the ionosphere. An increasing number of meteorologists, climatologists and hydrologists now are inclined to believe that in all probability it is necessary to take solar activity into account in the long-range forecasting of hydrometeorological conditions and climatic variations. The practical importance of the problems to be treated in this book, therefore, is obvious. One of the most important and at the same time one of the most poorly developed problems is the methods to be used in forecasting solar activity. At present we have only empirical-statistical methods for making such forecasts. It is essential to direct our efforts to the most attentive study of all conclusions concerning the nature of solar activity obtained using physicomathematical methods. The ultimate purpose is to raise the qualitative level of solar activity forecasting by a change from the unproductive empirical-statistical methods to the more promising physicomathematical methods.

However, the most significant obstacle in the solution of the sun-earth problem is the nature of the mechanisms involved in the relationship between solar activity and geophysical phenomena. A number of old concepts recently were found to be without basis as a result of the great successes in the investigation of the upper layers of the atmosphere, by meteorological rockets and artificial satellites, and the circumterrestrial part of interplanetary space, by means of space rockets. An acceptable theory of magnetic storms is again a rather acute problem.

In the new treatment of a number of problems associated with these phenomena, from the point of view of magnetohydrodynamics, allowance for the discovery of the earth's radiation belts and other circumstances has resulted in the Chapman-Ferraro theory, only recently considered rather promising, no longer corresponding to new data. The large number of hypotheses now proposed to explain these phenomena is evidence that this problem is rather far from solution. The discovery of solar X-radiation and the secondary radiation of the earth's atmosphere forces us to reconsider certain seemingly basic concepts concerning the nature of the ionosphere. However, the most acute problem still is an understanding of the mechanisms by which solar effects influence tropospheric processes. In contrast to geomagnetic and ionospheric processes for which more or less satisfactory physicomathematical theories have been proposed in the past, considering the then current level of factual information, no final theory has been advanced until now to explain solar effects on weather and climatic phenomena. It has been only recently that specialists in dynamic meteorology have made contributions in this direction.

The author feels that what has been said above should demonstrate that currently one of the most important aspects of the solar activity

investigation is the task of providing a theoretical basis for the enormous amount of empirical data. In certain cases this can be achieved by a critical study and comparison of proposed hypotheses against newly observed and experimental data. In other cases there is a need for new theoretical principles.

This monograph has six chapters. The first chapter is a systematic exposition of the current status of the problem of solar cycles. Attention is also given to such related problems as active longitudes, natural movements of sunspot groups along meridians and parallels, and displacements of the spot-forming zone during the solar cycle.

The second chapter considers a variety of problems associated with subphotospheric stratification and circulation on the sun. It is an introduction to a variety of ideas relating to the question of the structure of the internal, but at the same time, quite peripheral layers of the sun. The exposition of these problems is prefaced with an amount of material, to the extent considered necessary, relating to certain problems involved in hydrodynamics, thermodynamics and magnetohydrodynamics (e.g., information on convection theory, the theorems of circulation).

The third chapter is devoted to what might be called the "energy resources" of solar activity. The principal purpose of this chapter is to discover to what degree the energy of solar activity can be considered a part of the energy of circulation; or, on the other hand, whether it is not more rational to consider the second to be the product of the first. Of course, there also is a third possibility, that both are the result of the internal energy of the sun, i.e., in the long run the result of deep-seated thermonuclear reactions. Clarification of these problems is of basic importance, as it can lead to the solution of the problem of the internal or external origin of the energy of solar activity. At the same time, an attempt is made to estimate the energy of solar activity not only on the basis of information on the dynamics and magnetism of solar formations, but also by means of a quantitative estimate of active solar radiation.

The fourth chapter discusses problems associated with the effect of solar activity on the outer envelopes of celestial bodies, especially on the upper layers of the earth's atmosphere. This range of problems is extraordinarily broad and it was necessary to select only those aspects which were essential for our purpose. Therefore, considering the content of the fifth chapter, attention was concentrated on the problem of the heating of the upper layers of large bodies of the solar system, particularly the earth, by solar wave and corpuscular radiation.

The fifth chapter is an exposition of the current status of the problem of the manifestation of solar activity in the lower layers of the

earth's atmosphere, i.e., the sun-troposphere problem. In this aspect of the sun-earth problem, as in no other, the successful solution of a number of problems is dependent on the extent to which a critical evaluation is made of the numerous studies which have been made and continue to be made in this particular field. The attention given to review material in this chapter, therefore, should be somewhat greater than in the others.

Finally, the sixth chapter discusses manifestations of solar activity in the atmospheres of the major planets and comets.

Appendices include tables of certain solar activity indices which have been introduced somewhat recently. This information is scattered through many periodicals and it appeared desirable to summarize it in a single place.

CHAPTER 1
CYCLIC CHARACTER OF SOLAR ACTIVITY

Section 1. 11-Year Cycle

It is well-known that the cyclic character of solar activity is manifested most clearly in its 11-year cycle. Discussion of this problem should begin with a description of its principal properties.

The first vague references to a possible periodicity of solar phenomena, but without mention of any definite period, were made by Horrebow (Ref. 1). The observational data which he had at his disposal were scarcely sufficient for determination of the most probable period; this problem is not entirely clear because Horrebow's data were destroyed during the bombardment of Copenhagen by the British fleet. It is not impossible that there was a skeptical attitude toward Horrebow's opinion by such authorities as G. Cassini, Lemonnier, Lalande and Delambre, who were convinced of the absence of any regularity in the manifestations of sunspots (see Ref. 2) and, thereby, exerted an influence on this subject.

As is well-known, the discovery of the cyclic character of solar activity is credited to Schwabe. On the basis of his 20 years of observations he obtained a cycle with a duration of 10 years. After an exceedingly scrupulous investigation of archival data (only an insignificant part of which was published), R. Wolf devoted an entire series of articles to an extremely detailed study of the problem. He concluded that the cycle discovered by Schwabe was real. He found that a more precise value for its duration was 11.1 years. Wolf succeeded in obtaining more or less thorough data on sunspots for as far back as 1749. In addition, the epochs of maxima and minima of the 11-year cycles were traced as far back as the time of Galileo (1610). On the basis of these data, Wolf determined the value of the index which he had devised for each month, beginning with January, 1749. It should be noted, however, that daily data, taken separately, have a small weight even today. The mean monthly values since January, 1749, are quite representative. Still more fragmentary than the data dating to 1749 was that dating to 1700; Fritz determined the mean annual Wolf numbers R by using data back to this earlier date. The fragmentary character of the observational data for this period is already so great that the mean monthly values have an extremely low weight. The epochs of maxima and minima in the 11-year cycles are determined with an accuracy to one year from the

curves of the mean annual values R. However, smoothed Wolf numbers have been used since Wolf's time for a clearer determination of the cyclic character of solar activity: for the seventh month in a series of observed mean monthly values, the smoothed number is obtained by use of the formula

$$\bar{R}_7 = \frac{1}{2}\left(\frac{1}{12}\sum_{\mu=1}^{12} R_\mu + \frac{1}{12}\sum_{\mu=1}^{12} R_{\mu+1}\right). \qquad (1.1)$$

The mean monthly smoothed Wolf numbers determined in this manner are then used to determine the mean annual smoothed value W. It is assumed that in this way it is possible to determine the epoch of the maximum or minimum of solar activity in the 11-year cycle with an accuracy to one month, but this apparently is untrue. The fluctuating character of this activity with a natural fluctuation duration exceeding one month does not make it possible to assign any particular month to the extremum of the 11-year cycle. This is manifested with particular clarity in the current cycle, which is designated No. 19 in accordance with the system introduced at Zurich. A few comments should be made concerning this numbering. It is assumed that the epoch of the minimum of the 11-year cycle falling in the year 1755.2 is the epoch of the minimum of cycle No. 1. The years from 1749 to 1755 (the beginning of the latter) are assigned to the epoch of cycle No. 0 (obviously this cycle was not fully covered by systematic observations and has been represented only by the mean annual values determined by Fritz). Earlier cycles have been assigned negative numbers.

The cycle whose minimum, considered as the commencement, falls in 1954, has the designation, No. 19. An attempt to determine the epoch of the maximum of this cycle by use of formula (1.1) led to difficulties, forcing Waldmeier to be skeptical of the universality of this formula (Ref. 3). In general it can be assumed that smoothing by use of formula (1.1) has a double purpose: a) it is a means for avoiding the fluctuations, and, b) it is a method for eliminating the annual variation of Wolf numbers. A number of books contain a table of these maxima and minima of the 11-year cycles beginning with the year 1610.8 (Refs. 2, 4, 5). This table reveals that the length of a cycle can be determined by two methods, either as the interval of time between successive minima or as the interval between successive maxima, and is subject to considerable variations. In the first case, it can range from 8 to 15 years and in the second, from 7 to 17 years. There is a similar variation in the amplitudes of the cycles, that is, the differences of the heights of the cycle at minimum and maximum, and this is dependent on the difference of the heights of the maxima and the depths of the minima in different cycles. The deepest minimum had a value R = 0.0 and

the shallowest had a value R = 11.2; the lowest maximum had $R_M = 48.7$ and the highest $R_M = 189.9$. Here we have given the extremal monthly smoothed values, except for the last, which is an observed value.

This inconstancy of the length of the period and the different values of its amplitude justify the name "cycle" in contrast to the rigorous term "period". This will be discussed below in greater detail.

Figures 1-5 show the so-called cyclic curves for different indices of solar activity. The cyclic curve for any index is obtained in the following manner: the mean annual values of this index are arranged in accordance with the appropriate phases of the 11-year cycle, i.e., for all practical purposes in accordance with the corresponding years. Then, for comparability of different cycles the values are reduced to a single ordinate: the maximum value of the corresponding index is assigned the value 100 percent, and the values for the individual years are expressed in percent relative to the maximum. The curves for individual cycles then are superposed. In the diagrams accompanying the text, the epochs of minima have been superposed, but the superposition of epochs of maxima also is common practice.

Figure 1 shows curves of observed and smoothed Wolf numbers constructed in this way. Both these curves possess a maximum weight in comparison with all others because they were constructed from the maximum number of 11-year cycles. The curves in Figure 1 make possible the determination of the characteristics of the 11-year solar cycle with its steep and relatively short ascending branch, and longer and more gently sloping descending branch. Because certain cycles last more than 11 years on this diagram, and on those which follow, the 11th year of the cycle is shown, which in fact is the 12th, since the year of the minimum is assigned a zero. However, there are very few cycles with an 11th year, and the weight of the corresponding points, therefore, is insignificant.

Figure 2 shows curves of the areas of spots, faculas and prominences. It is shown that the cyclic variation of these three indices in general is identical. The areas of faculas cause some increase at the very end of the cycle, but this is noted only in the 11th year, and as already mentioned there are few such cycles and the weight of the corresponding point of curve 2 is small. A truer phenomenon is a less marked decrease in the area of prominences at the end of the 11-year cycle in comparison with spots and faculas. It is interesting that a similar situation also is observed during the development of individual active areas on the sun, i.e., of local phenomena having a far smaller characteristic time scale.

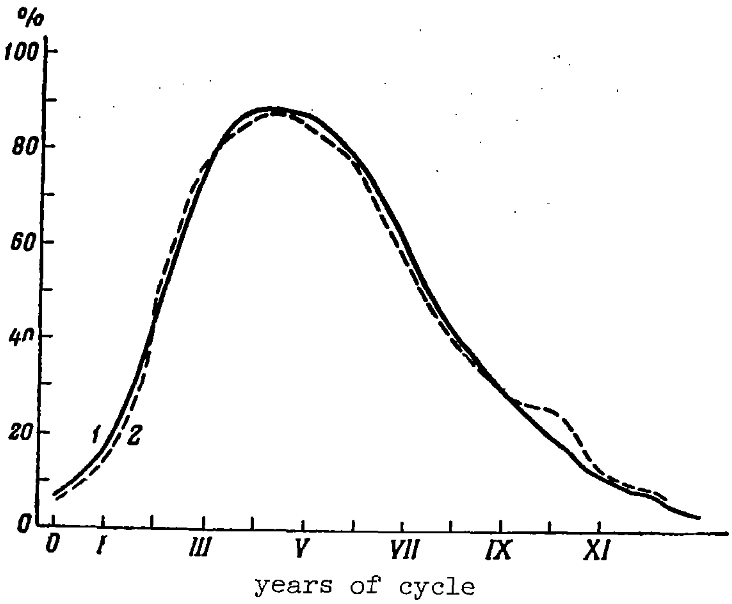


Figure 1. Cyclic curves of observed (1) and smoothed (2) Wolf numbers

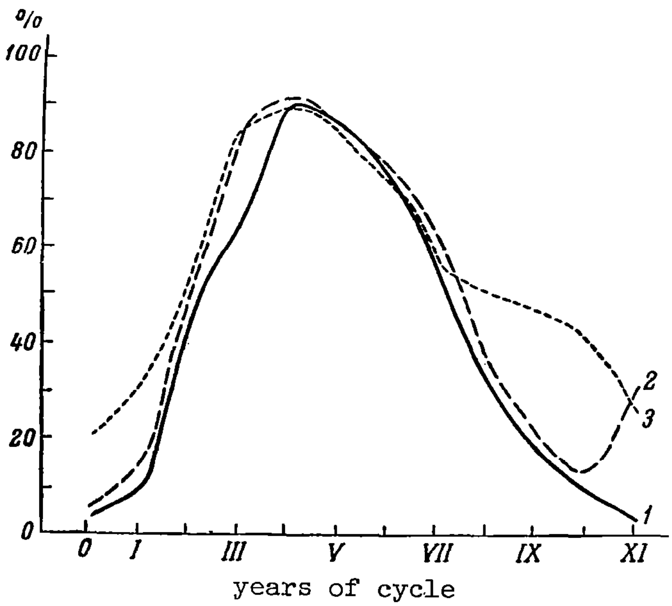


Figure 2. Cyclic curves of areas. 1-spots; 2-faculas; 3-prominences

Figure 3 shows the cyclic variation of the indices of duration and intensity of solar phenomena. It can be seen that the index $\bar{a}$ (curves 1 and 2)[1] reveals no appreciable variation in the 11-year cycle and its cyclic amplitude, that is, the difference of the values at maximum and minimum is insignificant; somewhat larger values are observed in the middle of the cycle (5th and 6th years) and somewhat smaller values at the beginning and end of the cycle. The curve of the index $\bar{a}$ not taking into account groups with a lifetime of only one solar rotation (curve 2, Figure 3), reveals some increase in values in the eighth and ninth years of the cycle but the cyclic amplitude nevertheless is small. The maximum group area (MGA) index (curve 3)[2] has a greater amplitude; it attains a maximum value in the epoch of the cyclic maximum R, i.e., in the fourth and fifth years of the cycle, and is characterized by a more rapid increase in the ascending branch and a slowing dropoff toward the end. However, the maximum of the 11-year cycle, according to this index, is rather gently sloping, and near the epoch of the maximum, according to the W index, the values of the MGA index in the preceding and following years do not differ too greatly from one another. Something similar is observed in the case of the average group area (AGA) index.[3] The cyclic variation of the different variants of this index is more or less identical, but it should be noted, nevertheless, that at the very beginning of the cycle (year of the minimum) the mean area of the group is less than near the maximum, or at the end of the cycle. The jump in the mean area from the last year of the preceding 11-year cycle to the epoch of the beginning of the new cycle is rather significant.

Figure 4 shows the cyclic variation of the Kopecky indices: F_0, the number of new formations; and, T_0, the lifetime of spot groups.[4] This diagram illustrates Kopecky's conclusion that there is an 11-year cyclic variation of the F_0 index, but no such variation of the T_0 index (Ref. 6). With respect to cyclic amplitude, the F_0 index can be compared only with Wolf numbers and spot areas. It is worth noting that the index attains a maximum in the fifth year of the cycle, whereas R and spot area have a maximum in the fourth year.

---

[1]See Appendix, Table I.
[2]See Appendix, Table III.
[3]See Appendix, Table II.
[4]See Appendix, Table IV.

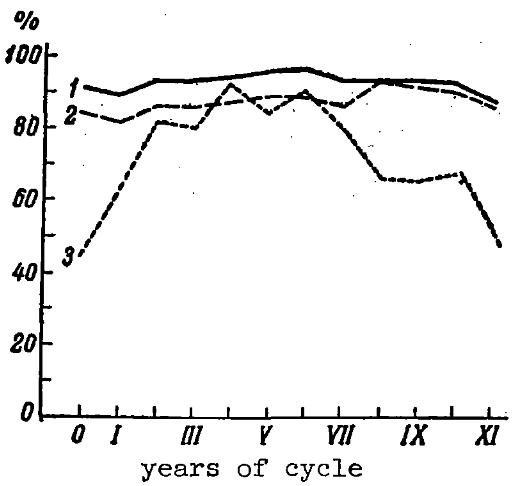


Figure 3. Cyclic variation of indices. 1-ā, with groups having a lifetime of only a single solar rotation taken into account; 2-ā, without taking into account groups having a lifetime of only a single solar rotation; 3-MGA index.

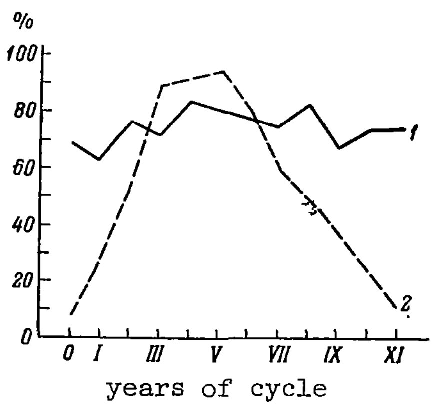


Figure 4. Cyclic variation of Kopecky indices. 1-T_0; 2-F_0.

Figure 5 shows cyclic curves of the Kleczek flare index and the Minnis and Bazzard ionospheric indices[1] and a curve for the F_2 layer. All three indices reflect the behavior of solar active wave radiation (Refs. 7-9), although of different components of this radiation. A certain similarity in their cyclic variation, therefore, is to be expected. The figure shows that such a similarity of the cyclic variation of the three indices is in fact observed. The cyclic amplitude of the F_2-layer index is greater than that for the E layer, which is also not unexpected, since in 1938 Appleton (Ref. 10) pointed out that simple critical frequencies give a greater cyclic amplitude for the F_2 layer than for the E layer. The behavior of the E- and F_2-layer indices is similar near the maximum of the cycle, but on the descending branch, the activity, determined using the F_2-layer index, drops off somewhat more rapidly than activity indicated by the E-layer index. The maximum value of the Kleczek index falls in the fifth year of the cycle, i.e.,

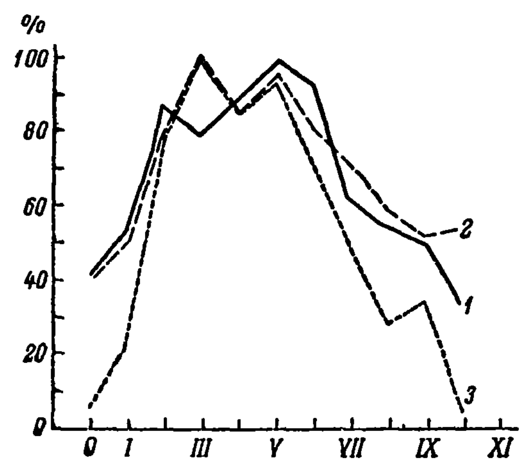


Figure 5. Cyclic variation of certain chromospheric and ionospheric indices. 1-Kleczek flare index; 2-Minnis and Bazzard ionospheric indices for E layer; 3-same for F_2 layer.

---

[1]See Appendix, Tables V and VI.

it lags somewhat behind the maximum according to Wolf numbers. This can be attributed to a certain predominance of large sunspot groups, in which flares are concentrated, at the very beginning of the descending branch of the cycle.

We now will proceed to the problem of internal relationships. As already noted, Figure 1 shows that the development of the 11-year cycle follows a slightly asymmetric curve; the ascending branch is slightly shorter and steeper than the descending branch. A similar result is obtained when superposing the cyclic curves (for individual cycles) of Wolf numbers for the epochs of their maxima. The interval of time between the minimum and maximum is called the time of increase of the cycle and is denoted T (or t). The interval of time between the maximum and the epoch when the smoothed mean monthly Wolf numbers attain values 7.5 is called the time of decline of the cycle and is denoted Θ (or τ). There are those 11-year cycles which are characterized by shallow minima; in such cases, the mean monthly Wolf number value may not decrease to 7.5 at the end of the preceding cycle. In the study of the behavior of the 11-year cycle, a frequently used index is the sum of Wolf numbers during the time of increase or the time of decrease of the cycle (usually using mean annual values), and also during the entire cycle (the corresponding notations are S_1, S_2, ΣR).

Waldmeier has established the following patterns in development of the 11-year cycle of solar activity (Ref. 11);

1) the higher the maximum, the shorter the ascending branch;

2) the higher the maximum, the longer the descending branch;

3) the higher the maximum, the stronger the spot-forming activity five years after the maximum;

4) the higher the maximum, the greater is S_2; and,

5) S_1 is almost not dependent on the height of the maximum.

If the following notations are introduced--R_M, maximum smoothed mean monthly Wolf number; R_5, smoothed mean monthly Wolf number five years after the maximum--then the relationships established by Waldmeier can be written as follows:

$$\log R_M = 2.58 - 0.14T, \qquad (1.2)$$

$$\Theta = 0.030R_M + 3.0, \qquad (1.3)$$

$$R_5 = 0.29R_M - 11.4, \quad (1.4)$$

$$S_2 = 40.6R_M - 572, \quad (1.5)$$

$$S_1 = 2538 + 0.4R_M. \quad (1.6)$$

With respect to formula (1.2) it should be noted that it was published by Waldmeier later than the others (Ref. 11). In Ref. 11, which gives the other relationships cited here, Waldmeier gave two formulas in place of (1.2); one of these formulas was for odd and one was for even cycles in the Zurich numbering system (see above). Waldmeier's initial formulas had the form

$$\log R_M = 2.48 - 0.10T, \text{ odd cycles}, \quad (1.7)$$

$$\log R_M = 2.69 - 0.17T, \text{ even cycles}. \quad (1.8)$$

These two formulas later were replaced by the single formula (1.2). According to Gleissberg, this was done because the long-sustained pattern in which the even cycles were relatively lower than the odd cycles was disrupted in cycle No. 18, which was higher than the preceding cycle, No. 17. This suggested that there was no great difference between even and odd cycles, as had been assumed earlier.

Of all the relationships (1.2)-(1.6), the clearest regression is (1.2). Attempts have been made to replace this logarithmic regression by a linear one. It should be mentioned that as early as 1931, (before the publication of Waldmeier's studies, Refs. 11-12), Ludendorf (Ref. 13) found a linear regression between the length of the ascending branch of the cycle and the excess of maximum R above the corresponding minimum. This dependence has the form

$$M - m = 4.43 + 3.15\left(1 - \frac{\Delta R}{100}\right), \quad (1.9)$$

where M is the epoch of the maximum of the cycle, m is the epoch of its minimum, and $\Delta R = R_M - R_m$, i.e., the difference between the mean annual Wolf numbers at the maximum and minimum. Xanthakis (Ref.14) has recently studied problems relating to the length of the ascending branch of the cycle and its height. V. F. Chistyakov (Ref. 15) also uses a linear regression between the length of the ascending branch of the cycle and the height of its maximum. In addition to the regressions (1.2)-(1.6), it should be noted that there are similar relationships between the mean annual values R in the descending branch of the cycle, and the mean annual value at the maximum (obtained by A. I. Ol', Ref. 16). Ol' obtained regressions between R in the first year after the

minimum, and in the remaining years of the ascending branch of the cycle. These regressions are of great importance in the prediction of solar activity (Ref. 17).

In 1949, S. M. Kozik noted (Ref. 18) that in the development of the first half of each 11-year cycle, i.e., along virtually the entire ascending branch, it is possible to detect two points of some significance. After the curve of development of the cycle has passed through these points, the angle coefficient of the tangent to the curve changes. This problem was later studied by Kopecky (Ref. 19), who confirmed Kozik's findings. Kozik and Kopecky established that the development of the 11-year cycle can be represented graphically by three straight lines of different slope. The first straight line begins from the epoch of the minimum and continues to a point at a mean distance of 13 months from the epoch of the cyclic maximum. This straight line is characterized by a rather considerable slope toward the time axis. The slope of the straight line decreases relative to the horizontal axis between the first and second breaks in the line. The third straight line corresponds to the descending branch of the cycle. The second "knee" in most cases coincides with the maximum of the cycle.

Kopecky established the presence of three types of 11-year cycles, differing in the position of the "knees" and the slope of the straight lines reflecting the cyclic development. At this point in the book, we cannot develop the concepts concerning the physical nature of the 11-year cycle; this will be done in the succeeding chapters. However, in order to understand the essence of the methods used for the representation of the cyclic curve, it is necessary to discuss briefly two of the principal working hypotheses.

From the time of discovery of the 11-year cycle, hypotheses have been proposed postulating an external, i.e., extrasolar, origin of this phenomenon. The not entirely regular periodicity of the phenomenon was attributed to either tidal effects induced by the planets, particularly Jupiter, or periodic encounters between the sun and meteor streams at the time of approach of the latter to their perihelion (Ref. 11).

The first of these hypotheses (Ref. 20) was the most fully developed and widely accepted. The rather close coincidence of the mean length of the 11-year solar cycle, 11.11 years, and the period of revolution of Jupiter, 11.86 years, suggested a tidal effect induced by that planet. The discrepancy of 0.75 year was considered attributable to the influence of other planets. If this hypothesis were true, the variations of solar activity obviously could be represented as the result of superposing of the influences of different planets, and since each of them has a constant period of revolution, the overall pattern of cyclic change could be represented in principle as the sum of individual harmonics. Over a period of many years, attempts have been made

to represent solar activity in just this way, i.e., as the result of the effect of an aggregate of individual periods associated with the times required for revolution of the major planets. The systematically obtained unsatisfactory results forced abandonment of this hypothesis.

Another solution was sought by Waldmeier in 1935 (Ref. 11). The principal argument in Waldmeier's working hypothesis can be summarized as follows. Each 11-year cycle should be considered a closed unit and the totality of 11-year cycles is, in fact, an alternation of enormous, similar explosion-like processes. It was on the basis of such a concept that Waldmeier derived those relationships mentioned above. However, it would be more correct to assume Waldmeier's point of view is as extreme as the hypothesis of a superposing of periods caused by external influences. It will be demonstrated that the 11-year cycles can in no way be considered as phenomena completely isolated from one another.

Prior to Waldmeier's investigations, when it was assumed that the cyclic character of solar activity could be represented in the form of a set of periods, a wide variety of mathematical algorithms were proposed for describing it. All constituted varieties of trigonometric series. For example, in Ref. 21, Kimura proposed the following formula:

$$R=\sum A\sin\left(\frac{2\pi}{T}t+H\right), \tag{1.10}$$

where A is amplitude, T is period, and H is phase. Kimura included 29 terms in this expression, each of which corresponded to a definite period. The period 11.11 years is of decisive importance. In Ref. 22, Oppenheim used a somewhat more refined approach; he introduced a periodic function with variable period θ:

$$0=\frac{2\pi}{T}t+\sum x_m\cos(m\psi t-\zeta_m), \tag{1.11}$$

where $\psi = \frac{2\pi}{T_1}$ (T_1 is another period of a higher rank), ζ_m is the variable phase (m = 1, 2, 3...), x_m is determined from the theory of Bessel functions as the argument of the corresponding function

$$\left.\begin{aligned}\cos\left(\frac{2\pi}{T}+x\sin\frac{2\pi}{T_1}t\right)&=\sum I_n(x)\cos\left(\frac{2\pi}{T}\pm n\frac{2\pi}{T_1}\right)t,\\ \sin\left(\frac{2\pi}{T}+x\sin\frac{2\pi}{T_1}t\right)&=\sum I_n(x)\sin\left(\frac{2\pi}{T}\pm n\frac{2\pi}{T_1}\right)t,\end{aligned}\right\} \tag{1.12}$$

where n = 1, 2, 3... ∞ .

If x sin $\frac{2\pi}{T_1}t$ is replaced on the left-hand side of the second of the equations, (1.12), by the expression

$$x_1 \cos\left(\frac{2\pi}{T_1} t - \zeta_1\right) + x_2 \cos\left(2\,\frac{2\pi}{T_1} t - \zeta_2\right) + \\ + x_3 \cos\left(3\,\frac{2\pi}{T_1} t - \zeta_3\right), \qquad (1.13)$$

the necessary formula for x is obtained. Finally

$$R = A_0 + A \cos \theta, \qquad (1.14)$$

where A_0 is the initial amplitude.

There also are various methods for representation of the 11-year cyclic curve based on the Waldmeier concept. One of the first attempts of this kind was the formula of Stewart, Panofsky and Eggleston (Ref. 23). This formula has the form

$$R = F\theta^a e^{-b\theta}. \qquad (1.15)$$

Here R is the mean monthly smoothed Wolf number (Ref. 2), F, a and b are constants dependent only on the number of the cycle, that is, changing from cycle to cycle, and e is the base of natural logarithms. Time θ is expressed in years and is determined from the beginning of the cycle (not to be confused with the length of the descending branch in formula (1.3)). The commencement of the cycle is not the epoch of the minimum of solar activity as indicated by Wolf numbers, but the epoch of appearance of the first groups of the new cycle; the two are by no means the same. Stewart and his coauthors have shown that F and b are dependent on a, and the constants in formula (1.15), therefore, are determined by a single parameter only. This problem was recently reconsidered by Yu. I. Vitinskiy (Ref. 17). Vitinskiy concluded that b is a nondependent parameter. The authors of formula (1.15) point out that it represents a true picture with an error not greater than 7 percent. The structure of the formula indicates that the cofactor θ^a characterizes the intensity of the cyclic explosion and the cofactor $e^{-b\theta}$ describes its attenuation. The intensity of the explosion and the rate of decline are dependent on the parameters a and b, respectively. F, however, has the value of a scale factor.

After reducing formula (1.15) to logarithmic form and differentiating, Gleissberg (Ref. 24) obtained:

$$\frac{d\ln R}{d\theta}=\frac{a}{\theta}-b; \tag{1.16}$$

but since

$$\frac{d\ln R}{d\theta}=\frac{1}{R}\cdot\frac{dR}{d\theta}, \tag{1.17}$$

then

$$\frac{dR}{d\theta}=a\frac{R}{\theta}-bR. \tag{1.18}$$

The left-hand side of (1.18) contains the annual change of the smoothed mean monthly Wolf numbers. The first term of the right-hand side characterizes the influence of new formations and the second term the dissipation of activity. It follows from (1.18) that the increase in Wolf numbers is directly dependent on the values of these numbers themselves, and inversely dependent on the time passing from the beginning of the cycle; a decrease in Wolf numbers is directly dependent on this number itself, and is not dependent on time. Gleissberg (Ref. 25) also proposed the following formula:

$$R=\frac{1}{2}R_M\left[1+\frac{3}{2}\cdot\frac{t-t_H}{t_M-t_H}-\frac{1}{2}\left(\frac{t-t_M}{t_M-t_H}\right)^3\right]. \tag{1.19}$$

Here, R is the mean annual Wolf number, smoothed for three successive values, t is the epoch for which R is determined, t_M is the epoch of the maximum, R_M is the Wolf number at the maximum, and t_H is the time when $R = \frac{R_M}{2}$.

The representation of observations by use of this formula gives a probable error of the mean annual value R of about 4 units. Chvojkova (Ref. 26) has proposed the following formula:

$$R=\frac{R_M}{2}\left[1-\cos\frac{2\pi t}{aT+(1-a)t}\right], \tag{1.20}$$

where t is the time read from the beginning of the scale, and T is the duration of the cycle,

$$a=\frac{\frac{T_1}{T}}{1-\frac{T_1}{T}}. \tag{1.21}$$

Here, T_1 is the duration of the ascending branch of the cycle. As pointed out by A. I. Ol', formula (1.20) has no advantage over formula (1.15) (Ref. 16).

In summary it can be said that neither the method of superposing epochs nor the Waldmeier method in unmodified form properly reflects reality. Despite a certain independence of each individual 11-year cycle, it is impossible to consider the individual cycles to be unrelated to one another. The representation of the 11-year cycle by use of a formula similar to the Stewart-Panofsky formula should be combined with a periodogram representation of the alternation of cycles. However, this has not yet been done.

Section 2. 22-Year Cycle

The first claims of the presence of a 22 to 23-year cycle of solar activity were made by R. Wolf (Ref. 27). In 1913, the existence of such a cycle was noted by Turner (Ref. 28). Turner felt that one 23-year cycle consisted of two 11.5-year cycles (this length of the 11-year cycle corresponded closer to reality in Turner's opinion than the duration of 11.11 years obtained by other investigators) and the two halves of this cycle were asymmetric. Turner assumed that the first half of a "period" is of greater amplitude and somewhat greater inclination (of the ascending branch) than the second. The decisive confirmation of the physical reality of the 22 to 23-year cycle was the well-known discovery of Hale, et al. (Ref. 29). By studying the signs of magnetic polarity of approximately 2,000 groups for the years 1908-1925, they established that during three 11.5-year cycles only 41 groups did not conform to the rule of change of sign of magnetic polarity with a transition to the next cycle. The discovery of the 22-year cycle as a phenomenon reflecting important features of the physics of solar activity attracted great attention to this cycle. Turner (Ref. 28) demonstrated that the odd 11-year cycles, No. 13 (1888-1899) and No. 15 (1912-1923), differ from the even cycles No. 12 (1876-1887) and No.14 (1900-1911) in the following characteristics: the spots in both hemispheres were an average of 1° farther from the equator in odd cycles than in even cycles; in odd cycles the spot area was greater than in even cycles, with the ratio of these areas being 143:126. As in formula (1.15), Turner uses as the beginning of a cycle, not the epoch of the minimum, as indicated by Wolf numbers, but rather the appearance of spots in the new cycle, i.e., spots having a magnetic polarity characteristic of the spot groups of the new cycle. The alternation of the heights of the 11-year cycles noted by Turner for four such cycles is observed rather stably (see Table 1).

Table 1

No. of cycle	1	2	3	4	5	6	7	8	9	10	11
R...........	86	106	154	132	48	46	71	138	124	96	139

No. of cycle	12	13	14	15	16	17	18	19
R...........	64	85	64	104	78	114	152	190

As shown in this table, the stable alternation of the heights of the cycles follows the rule, even-low, odd-high, and this condition has been satisfied since cycle No. 10. This pattern did not prevail in earlier epochs. Turner assumed that this was because of the inadequate accuracy of sunspot observations in the 18th and early in the 19th centuries. However, when there was a departure from the rule, even-low and odd-high, in the transition from cycle No. 17 to cycle No. 18, i.e., in our time, this explanation had to be rejected. There is basis for assuming that the disruption in the pattern of alternation of the heights of the cycles both in the late 18th century and beginning of the 19th century, and in our time can be attributed to the same factor, associated with the long-term variation of solar activity.

As indicated by Kopecky (Ref. 6), it is interesting to note that the F_0 index, i.e., the number of new formations in the epoch of the maximum, as well as the total value of this index for the 11-year cycle, continue to reflect the even-low and odd-high pattern even with the transition from cycle No. 17 to cycle No. 18. Figure 6 is a graphic representation of the variation of the cyclic sums of Wolf numbers and of the F_0 index. The numbers of the 11-year cycles have been plotted along the x-axis and the sums ΣW and ΣF_0 (for convenience, the latter have been reduced by a factor of 10) have been plotted along the y-axis.

This figure shows clearly how the alternation of the heights of the cycles is dependent on their evenness or oddness. Using the Wolf numbers, it is found that cycle No. 18 was higher than No. 17, i.e., the pattern is disrupted; whereas, using the $\Sigma F_0/10$ index no such disruption is observed, although according to Kopecky the data on F_0 for cycle No. 18 must be considered to be preliminary.

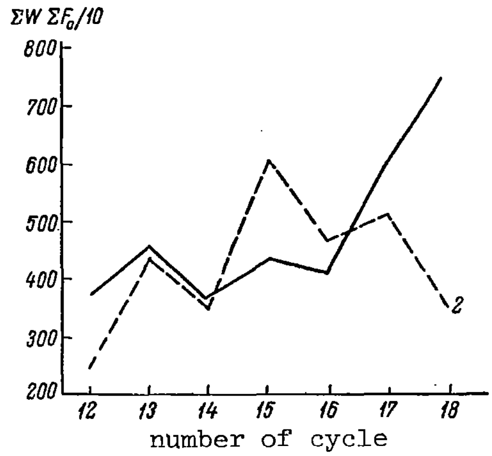


Figure 6. Variation of sums of Wolf numbers (1) and F_0 index (2) for cycle

The amplitude of the 22-year cycles, that is, the difference in the heights of the maxima of the high and low cycles of a pair, is different for different 22-year pairs. By the amplitude of an 11-year cycle, we mean the difference between the mean annual values R at the maximum and minimum. Table 2 gives the numbers of the 11-year cycles, the mean annual observed Wolf numbers in the years of the maximum and minimum, the amplitude of the 11-year cycles, the amplitude of the 22-year cycles and the ratio of the second to the first for even and odd cycles.

This table gives only those pairs of 11-year cycles where there is no question of the presence of a 22-year cycle; in a pair of 11-year cycles there is a low-even and high-odd pattern. The table shows that the ratio of amplitudes in 22- and 11-year cycles is greater in the case of an even cycle and lesser in the case of an odd cycle; this can be attributed to the greater height of the odd 11-year cycles.

It is difficult to estimate the amplitudes of the 22-year cycles by use of other indices than Wolf numbers because of the limited availability of series of all other characteristics of solar activity and because the 22-year cycle, as we have seen, does not show up when certain other indices are used. Here we will cite the amplitudes of the 22- and 11-year cycles using the index of corpuscular activity δP, and the index of spot area (Table 3).

The problem of the rule for combining 11-year cycles into 22-year cycles was considered by M. N. Gnevyshev and A. I. Ol' (Ref. 30). This problem should be formulated as follows: does a 22-year pair begin with

Table 2

No. of 11-year cycle	Observed mean annual extrema		Amplitude of 11-year cycle	Amplitude of 22-year cycle	Ratio for even 11-year cycle of pair	Ratio for odd 11-year cycle of pair
	max.	min.				
2	106	11	95	48	0.505	0.326
3	154	7	147			
6	46	0	46	23	0.500	0.333
7	71	2	69			
10	96	4	92	43	0.468	0.326
11	139	7	132			
12	64	3	61	21	0.345	0.266
13	85	6	79			
14	64	3	61	40	0.656	0.389
15	104	1	103			
16	78	6	72	36	0.500	0.334
17	114	6	108			
				mean	0.495	0.330

an even or an odd cycle? It is obvious that such pairs can be formed in two ways: a) combining of an odd cycle with the succeeding even cycle; and, b) combining an even cycle with the succeeding odd cycle.

Gnevyshev and Ol' computed the correlation coefficients between ΣR for the cycle in each of these two combination methods. It was found that if an odd cycle is used as the beginning of a 22-year cycle, the correlation coefficient is $+50 \pm 0.24$. However, if it is assumed that the 22-year cycle begins with an even 11-year cycle, the correlation coefficient is $+0.91 \pm 0.11$. This result undoubtedly shows that a 22-year cycle begins with an even 11-year cycle.

Among all the nineteen 11-year cycles considered by Gnevyshev and Ol'--from No. 2 to No. 17--there were only two cycles, No. 4 and No. 5, which fail to conform to a direct regression on the correlation diagram. Conclusions concerning the relationships existing between 11-year cycles constituting a 22-year pair can be drawn in particular from a study by V. F. Chistyakov (Ref. 31), who was able to note the following two facts:

1. The total intensity of the 11-year cycles making up a 22-year pair is a linear function of the duration of the 22-year cycle. The

corresponding regression equation has the form

$$R_\sigma = 748.8 - 26.748\tau, \tag{1.22}$$

where R_σ is the total intensity of the even and odd cycles constituting the 22-year pair, and τ is the total duration of the double cycle.

This relation applies for the following pairs of 11-year cycles: Nos. 0-1, 2-3, 6-7, 12-13, 14-15 and 16-17. However, cycles Nos. 4-5, 8-9 and 10-11 fall along a straight line with the equation

$$R_\sigma = 717.5 - 20.856\tau. \tag{1.23}$$

2. On the basis of the relationship established by Gnevyshev and Ol', Chistyakov obtained the following regression, correct for these pairs of 11-year cycles: Nos. 2-3, 6-7, 10-11, 12-13, 14-15 and 16-17.

This regression has the form

$$R_{r_{n+1}} = 4.2 + 1.398R_{r_n}. \tag{1.24}$$

The results cited above indicate a quite close relationship between the 11-year cycles forming the 22-year pair. This is a convincing argument against the initial Waldmeier concept that each 11-year cycle is completely independent of the others. At the same time, certain epochs are characterized by a disruption of the usual relations in the 22-year cycle; as already mentioned, this is associated with the existence of a cycle of long duration. We will now proceed to a description of the characteristics of that cycle.

Section 3. Long-Term (80-90-Year) Cycle

We will begin with problems of terminology; in addition to the names used in the above heading, other names assigned to this cycle are: 80-year period, 7-8 or 8-9 cycle period, long cycle, long-range cycle and secular cycle.

Wolf mentioned (Ref. 32) that in addition to the 11-year (and 22-year) cycles there is a quasi-periodic variation of great duration. It was initially assumed that this cycle was 55.5 years long, but later he became inclined to the opinion that the duration was between 66-2/3 and 83-1/3 years. As pointed out by Waldmeier (Ref. 33), Wolf, who had at his disposal data for only the two long cycles of the 18th and 19th centuries, with the data for the latter still incomplete, should be considered the discoverer of this cycle, much like Schwabe, who discovered

the 11-year cycle, also having at his disposal data for only two such cycles. In 1909 Wolf called attention to the long-term cycle in Ref. 34. Later this cycle was studied by Gleissberg (Ref. 2), Eygenson (Refs. 35, 36), Stetson (Ref. 37), Dizer (Ref. 38), Balli (Ref. 39), Waldmeier (Ref. 40), and others.

We will first discuss the methods used for detection of the long-term cycle. As already noted by Wolf, and as pointed out by Gleissberg (Ref. 2), the 80-year cycle is primarily a quasi-periodic change of the heights of the maxima of the 11-year cycles. But as mentioned in Section 1 of this chapter, the remaining characteristics of the 11-year cycle also are related to the height of the maximum. Therefore, the 11-year cycle to a certain degree also reflects the quasi-periodic character of the change of other cyclic parameters.

One of the principal methods that can be used for detecting the 80-year cycle in a sequence of 11-year cycles is secular smoothing, proposed by Gleissberg (Refs. 2, 4). This method involves the forming of a moving average from four successive values, usually either the heights of the 11-year cycles at their maxima or the corresponding values at the minima. Two such moving averages are used, the second being formed of terms displaced by one counting unit, that is, the second moving average

Table 3

No. of 11-year cycle	Extrema for areas, in millionths of hemisphere		Extrema for δP		Amplitudes of 11-year cycles		Amplitudes of 22-year cycles	
	max.	min.	max.	min.	area, in millionths of hemisphere	δP	area, in millionths of hemisphere	δP
12	1155	24	27	-32	1131	59	309	19
13	1464	78	46	-35	1386	81		
14	1191	29	13	-46	1162	59	346	26
15	1537	7	39	-48	1530	87		
16	1390	55	33	-22	1335	55	684	23
17	2074	88	56	-21	1986	77		

does not involve the first term, but a fifth is added. A simple mean is taken from these two moving averages, relating to the epoch of the mean of the five used. For example, if there are five successive values, R_M, that is, the values for 5 epochs of maxima of 11-year cycles, secular smoothing gives

$$\bar{R}_M = \left(\frac{R_M^{(1)} + R_M^{(2)} + R_M^{(3)} + R_M^{(4)}}{4} + \frac{R_M^{(2)} + R_M^{(3)} + R_M^{(4)} + R_M^{(5)}}{4} \right). \tag{1.25}$$

R_M refers to the epoch which is the simple mean of five epochs of maxima of cycles entering into the formula. Gleissberg used this method for smoothing epochs of minima and maxima of 11-year cycles beginning in the 18th century and the heights of the maxima and the values at the minima of such cycles beginning with 1705. We note that $R_M^{(i)}$ is the maximum mean monthly value of the Wolf number for a cycle smoothed using formula (1.1). Thus, the smoothed monthly values $R_m^{(i)}$, the lowest in the cycles, are used as the initial values also in secular smoothing of cyclic minima. Figure 7 shows the variation of the maxima of 11-year cycles as smoothed by Gleissberg; the numbers of the cycles are plotted along the x-axis (we note again that the 0 cycle is that having the epoch of the minimum in 1745); smoothed Wolf numbers are plotted along the y-axis. In his earlier studies (Refs. 42-44), Gleissberg still did not use the method of secular smoothing in the form which is given by formula (1.25), but used a simple, moving average of four successive values of Wolf numbers at the maxima of 11-year cycles. The results obtained using that method differ little from those obtained using formula (1.25). Figure 8 shows that using this method, sometimes called the four-maxima method, the long-term cycle in the 19th century is slightly higher than the similar cycle in the 18th century. This difference is not noted when there is secular smoothing using formula (1.25).

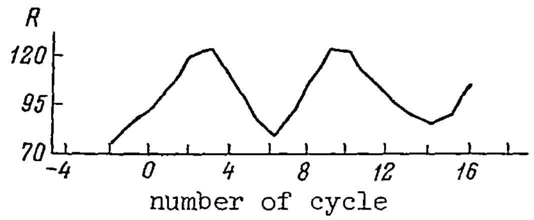


Figure 7. Variation of maxima of 11-year cycles, smoothed by Gleissberg method

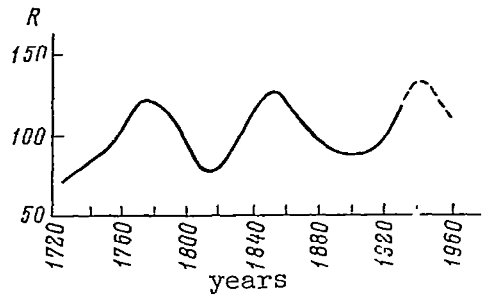


Figure 8. Variation of maxima of 11-year cycles, smoothed by the four-maxima method

Another method for detecting the long-term solar cycle is based on the forming of moving averages or moving sums. As stated in Ref. 45, if in some initial series there is a periodic term with the period P, in a series formed from moving averages of the values of the first of m terms, the amplitude of the period P decreases in the ratio

$$K=\frac{\sin m\frac{\pi}{P}}{m\sin\frac{\pi}{P}}. \tag{1.26}$$

However, if $P>m$, the larger the value m, the smaller will be the amplitude of the period P. On this basis it is possible to form moving sums of the mean annual values of the Wolf numbers for the entire available series of years, i.e., with sufficient reliability for 11-year periods since 1749. In this method the 11-year cycle would be excluded entirely if it was a rigorous period and not a cycle. Since this does not occur, certain traces of an 11-year cycle should remain even after the formation of the moving sums. In addition, the formation of moving sums using 11 values does not exclude a 22-year cycle, but only decreases its amplitude somewhat. In applying the moving sums method for the purpose of obtaining the long-term cycle it, therefore, is more rational to obtain such sums using 23 rather than 11 terms. Such a method should exclude or at least greatly attenuate the 11- and 22-year components. An odd number of terms is convenient in the sense that it makes it possible to determine precisely the year to which the value of the moving sum applies; the moving sum obviously differs from the moving average only in scale. Such a moving summing of a series of Wolf numbers was done in Ref. 46, and the result is shown in Figure 9. The irregularity of the curve can be attributed to remnants of the 11- and 22-year cycles. Using such a method for the detection of the long-term cycle its amplitude in the 18th century is higher than in the 19th century.

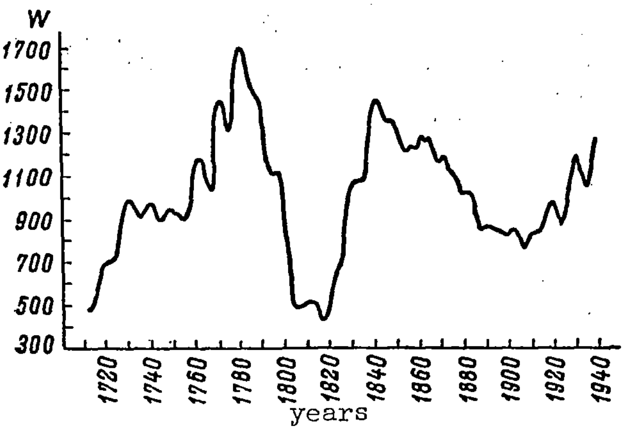


Figure 9. Curve of solar activity using a 23-year moving mean for period 1700 = 1947

We should mention still another elementary method for detection of the long-term cycle, proposed by M. S. Eygenson (Ref. 47). It involves the paired summing either of the sums of R for the 11-year cycle or R_M for this same interval. Taking of the sum R or R_M excludes the 11-year cycle because each cycle is represented by a single value only; paired summing excludes the 22-year cycle. The result of summing is given in Table 4.

Table 4

No. of 11-year cycle	22-year cycle	ΣR_M
1-2	1755-1774	192
3-4	1775-1797	286
5-6	1798-1822	93
7-8	1823-1842	209
9-10	1843-1866	220
11-12	1867-1888	203
13-14	1889-1912	149
15-16	1913-1932	182

The last column gives the sum R_M for the two 11-year cycles forming the 22-year cycle. In this table, giving the combination of 11-year cycles, an odd cycle is combined with the even cycle following, which violates the Gnevyshev-Ol' rule formulated in Ref. 30. Eygenson also gives a variant in which an even 11-year cycle is combined with the following odd cycle. Applying this method of grouping of 11-year cycles to earlier cycles, we obtain the results given in Table 5.

We will now discuss still another method for detecting the long-term solar cycle. This method involves the construction of integral curves, and is used rather widely in geophysics (Ref. 49). The first attempt to apply integral curves to the study of solar activity was made by P. P. Predtechenskiy and B. S. Gurevich (Ref. 50). This method was used later by A. A. Girs and I. M. Soskin (Refs. 51, 52). The latter author obtained the integral curve of Wolf numbers shown in Figure 10. The integral curve makes it possible to detect peculiarities which would not be detected using ordinary moving summing. For example, it clearly follows from a study of Figure 10 that the long-term cycle of the 18th century is higher and steeper than the same cycle in the 19th century. However, as noted above, no firm conclusions can be drawn on this basis at present because the two smoothing methods have yielded opposite results. Ringnes recently published (Ref. 53) a communication concerning the behavior of groups with a lifetime of only one solar rotation in the long-term solar cycle. He found that during the period from 1880 through 1956 the maximum percentage of such groups was observed at the beginning of the current century, i.e., in an epoch of the minimum of the long-term cycle. Taking into account that groups with a lifetime of only a single solar rotation usually have a small area and recalling that according to the MGA index the minimum of the long-term cycle fell at the beginning of the current century, it is possible to understand why the maximum number of groups with a lifetime of only a single solar rotation also fell in these same years.

The long-term variations of the strength of sunspot magnetic fields is of considerable interest. In the 11-year cycle there is no observable dependence of mean field strength on the phase of the cycle. In fact, there is a relationship between the area of sunspot groups and the strength of their magnetic fields only for spots having an area up to 300 millionths of the hemisphere; in this case, an increase of area is accompanied by an increase of field strength, reaching approximately 2,700 Oe when S = 300 millionths of the hemisphere (Ref. 4). Virtually no such relationship is observable in the case of spots of larger dimensions. If it also is taken into account that the AGA index is not too sharply expressed in the 11-year cycle, it becomes understandable that the mean strength of the magnetic field of spots is virtually constant in the 11-year cycle. This is not true for the long-term cycle. It is true that at present the limited availability of data does

Table 5

No. of 11-year cycle	22-year cycle	ΣR_M
0-1	1745-1765	169
2-3	1766-1783	260
4-5	1784-1809	180
6-7	1810-1832	117
8-9	1833-1855	262
10-11	1856-1877	235
12-13	1878-1900	149
14-15	1901-1922	168
16-17	1923-1943	192

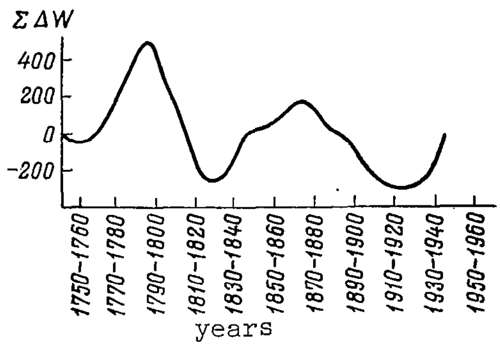


Figure 10. Wolf number integral difference curve

not make it possible to draw more or less final conclusions, but it can be noted that there is a "secular" (it would be more correct to say "long-term") variation of the magnetic field strength of spots (Ref. 54). Bell established that the number of groups with large strengths in epochs of maxima of the different 11-year cycles does not remain constant. In particular, the maximum of the current cycle of solar activity (No. 19), falling in 1957-1958, differed in that being extremely intense on the basis of the Wolf number index, at the same time it was the lowest in many cycles with respect to the number of groups with large magnetic fields (Ref. 54). Previously the mean field strength of groups on the ascending branch of the current long-term cycle had decreased somewhat, but this could be the result of nonuniformity in the

series of observations, since the progress of spectroscopic techniques has made it possible to record increasingly weaker magnetic fields. However, the observed phenomenon apparently cannot be attributed to this single effect.

A comparison of determinations of the magnetic flux in spots and their groups for different years, therefore, is of interest. An appropriate table is available in an article by Kiepenheuer in the book "The Sun" (Ref. 4), edited by Kuiper. This table shows that the flux of both the leading and the following spots of bipolar groups in both the northern and southern solar hemispheres attained maximum values in cycle No. 15, specifically, in 1917-1919. In cycle No. 18, which considerably exceeded No. 15 in intensity, all flux values were appreciably lower.

Similarly, there was a decrease of the values of the differences between the fluxes of the leading and following spots of bipolar groups of both the northern and southern solar hemispheres. Thus, the presence of a long-term cycle in variations of the magnetic characteristics of sunspots, if still not demonstrated, in any case is very probable.

We will now proceed to the problem of the duration of the long-term cycle of solar activity and its stability. When making his secular smoothing in Ref. 55, Gleissberg found that the length of this cycle averages 78 years, causing the cycle to be called an 80-year cycle. This result was obtained by Gleissberg using the epochs of the maxima of the 11-year cycles, making it possible to extend the investigation to the 17th century, which would have remained unstudied if Gleissberg had used R values.

In 1955 Gleissberg made an attempt to use far earlier 11-year cycles, established on the basis of information concerning the frequency of auroras (Ref. 55). Gleissberg assembled data beginning with 290 A.D. Thus, by applying secular smoothing to years of maximum frequency of auroras, assumed to be years of solar activity maxima, Gleissberg was able to obtain data concerning seventeen 80-year cycles. According to the nomenclature adopted by Gleissberg, the long-term cycle being experienced at the present time is No. 2; the cycle with a maximum in the last centruy was No. 1; and, the cycle with a maximum in the 18th century was No. 0.

Earlier long-term cycles apparently should be assigned negative numbers. It follows from the data obtained by Gleissberg (Table 6) that during the last 16 centuries the length of the long-term cycle has varied in the range of 5 to 11 11-year cycles. The mean length during this time was 7.1±3.3 11-year cycles. If we assume, as Gleissberg did, that the mean length of the 11-year cycle is 11.1 years, the mean length

Table 6

No. of 80-year cycle	No. of 11-year cycle	Length of 80-year cycle (in 11-year cycles)
-17	-126	8
-16	-118	7
-15	-111	6
-14	-105	8
-13	- 97	6
-12	- 91	9
-11	- 82	7
-10	- 75	7
- 9	- 68	11
- 8	- 57	6
- 7	- 51	5
- 6	- 46	6
- 5	- 40	7
- 4	- 33	7
- 3	- 26	7
- 2	- 19	9
- 1	- 10	6
0	- 4	6
1	2	7
2	9	-

Note. The following cycles are considered negative: 11-year cycles before 1745, and 80-year cycles before the 18th century

of the 80-year cycle should be 78.8 years. The numbering of the 11-year cycles in Table 6 has been adjusted to correspond to the long-term cycles listed in the first column.

The conclusions drawn by Gleissberg illustrate the variations of the length of the long-term cycle, but reveal nothing of the variations of its amplitude. I. V. Maksimov, in discovering an 80-year cycle in variations of the width of growth rings in sequoias, reported in Ref. 56, found that the amplitude of this cycle, insofar as can be judged from this index, changes during the period from 1925 B. C. through 1905 A.D. by a factor as great as 2. The long-term cycle, like the 11-year cycle, therefore, experiences both variations in length and variations in amplitude. In order to explore this problem further and also

clarify the relationships between the intensity of the long-term cycle and the lengths of its ascending and descending branches, a study was made on the basis of the same data used by Gleissberg in the reference cited above, i.e., data on the 11-year cycles in the historical past, primarily on the basis of information on auroras collected by Schove (Ref. 57).

It is true that Gleissberg, who apparently used the first variant of Schove's summary, did not have at his disposal the qualitative characteristic of the intensity of cycles which Schove cites in Ref. 57, and possibly did not ascribe importance to these data, carrying out his secular smoothing solely for the epochs of the 11-year cycles, not their amplitudes. More or less reliable data on the intensity of the 11-year cycles are available in Schove's study beginning with the cycle whose maximum fell in 214 B.C., and ending with the current cycle No. 19 (according to the Zurich classification).

It should be noted that although the Schove index (auroras) undoubtedly is better as a characteristic of solar activity than the data used by I. V. Maksimov, for a number of reasons this characteristic can result in considerable deviations in the determination of the epoch of the maximum and minimum of the 11-year cycle (there is an opinion that such a deviation can attain half an 11-year cycle). However, such a possible error should not be reflected in the determination of the epochs of long-term cycles; this applies to the same degree also to the amplitude of the cycle, which, in general, should be represented quite well by the auroral indices. The epochs of the maxima of the 11-year cycles contained in the Schove summary, as well as the intensities of these cycles in a 9-unit scale, were smoothed using formulas for secular smoothing (formula (1.25), and a similar formula for epochs). Since there are no data on intensity for cycles falling between 203 and 213 A.D., and 272 and 283 A.D., it was necessary to begin the series with the epoch falling near the beginning of the fourth century A.D.

The following facts were established from an investigation of data subjected to secular smoothing:

1. Over the course of almost two thousand years, variations of solar activity with a duration of approximately 80 years have in certain cases been observed quite clearly, such as in the 16th, 14th and 11th-12th centuries, and at other times have been expressed considerably less distinctly, such as in the 8th-9th and 5th-6th centuries; at times, such as in the 17th century, the phenomenon was virtually undetectable.

2. Rather frequently "90-year cycles" are of different durations, therefore corresponding to Gleissberg's conclusions. Tables 7 and 8 give data for distinctly (Table 7) and less distinctly (Table 8) expressed long-term cycles. These tables show that the mean length of

Table 7

Range of years	Ascending branch, years	Descending branch, years	Total length, years	Ratio of intensity at maximum to depth at minimum
1803-1894	40	52	92	2.4
1672-1802	77	53	130	1.4
1470-1603	80	55	135	1.6
1315-1414	46	52	96	2.8
1027-1149	93	29	122	2.0
885-1026	53	79	132	2.8
321-397	39	38	77	3.0

Table 8

Range of years	Ascending branch, years	Descending branch, years	Total length, years	Ratio of intensity at maximum to depth at minimum
1193-1228	10	22	32	1.1
787-884	53	45	98	1.3
733-787	12	42	54	1.1
520-606	45	41	86	1.1
452-520	47	18	65	1.1
4-95	56	37	93	1.2

distinctly expressed cycles exceeds 100 years, and the length of less distinctly expressed cycles is as great as 70 years. According to our data, the general mean length is approximately 93 years, which is greater than obtained by Gleissberg (79 years) and Maksimov (84 years). We avoid giving such precise values as those given by Gleissberg and Maksimov because we consider accuracy to one year, and especially to 0.1 year, is completely fictitious. The following is a possible explanation of why our mean values, on the basis of the same data, exceed those obtained by Gleissberg. Although Gleissberg proceeded on the basis of secular smoothing of the epochs of the maxima of 11-year cycles, as the duration of the long-term cycle he used the interval of time from one minimum to the next, whereas we considered the long-term cycle as the interval from one maximum to the next.

It was shown that even the 11-year cycles do not have the same duration if, in the one case, the interval between two maxima is considered the duration of the cycle, and, in the other case, the duration is computed as the interval between two successive minima, since in the first case the dispersion of lengths will be somewhat greater. This should be even more true of the long-term cycle. It is impossible to overlook the fact that the more clearly expressed long-term cycles are somewhat longer than the poorly expressed cycles, although at the same time the first are the more intense. Thus, if there is a certain relationship between the length of a long-term cycle and its intensity, this dependence is direct rather than inverse, whereas the opposite obtains for the 11-year cycle.

3. The last columns of Tables 7 and 8 make it possible to compare the cyclic ratios (i.e., the ratios of the heights at the maximum and minimum) for long-term cycles with the same ratio for 11-year cycles, using the mean values of these ratios in both cases. For 80-year cycles this ratio does not exceed 3:1, whereas, for 11-year cycles it is of the order of 20:1.

4. The correlation coefficients between the lengths of the ascending branches of long-term cycles and their intensity have values +0.60 for clearly expressed cycles (7 pairs) and +0.58 for all cycles (13 pairs). No relationship was discovered between the intensity of the long-term cycle and the length of its descending branch.

The most important conclusion is that there is a direct, although not very distinctly expressed relationship between the length of the ascending branch of the long-term cycle and its intensity at the maximum. On the other hand, it has been seen that the 11-year cycle is characterized by an inverse dependence between the length of the ascending branch and the height of the cycle. This conclusion is significant in that it indicates an important difference between the 11-year and long-term cycles (in addition to the trivial difference in time scale). This, as well as the very small value of cyclic amplitude in comparison with the 11-year cycle, indicates that there is an important physical difference between the 11-year and long-term cycles.

We will now cite certain additional information concerning change in the intensity of solar phenomena and related geophysical processes during the long-term cycle of solar activity. By comparing the numbers of groups making one and two appearances during the 11-year cycles, Nos. 13-16, Eygenson noted that the ratio of the number of repeating groups to the number of nonrepeating groups increases monotonically from the 13th to the 16th cycles. The increase in solar activity in at least the initial stage of the current 80-year cycle, therefore, was caused largely by repeating groups which usually are larger than groups existing for only one solar rotation.

Then, by introducing three gradations of group areas (from 200 to 500, from 500 to 1,000, and >1,000 millionths of the hemisphere), Eygenson demonstrated that in the 11-year cycle, No. 16, as compared with cycle No. 14, the number of groups falling in the first area gradation increased by 12 percent, in the second by 35 percent, and in the third by 290 percent. The cyclic ratio in the 80-year cycle, therefore, increases with a transition from the lesser to the stronger manifestations of solar activity (Refs. 58, 59).

As noted in Ref. 47, a qualitatively similar picture also can be observed in the 11-year cycle. At the same time, the total number of spot groups on the ascending branch of the long-term solar cycle can even decrease somewhat (Ref. 58). Thus, with an approach to the maximum of the 80-year cycle, the "absolute number of less intense phenomena decreases, and the absolute number of more powerful phenomena increases." However, the picture was somewhat different directly at the maximum of the current long-term cycle; for example, there was no large number of very large groups, so that this problem apparently is more complex than believed only recently.

On the ascending branch of the current long-term cycle, there also was noted a relative increase in the number of strongly expressed geophysical phenomena caused by solar activity. For example, the percentage of strong magnetic storms in the 11-year cycle No. 14 was 19, but in cycle No. 17 it had increased to 54.

In Refs. 60 and 61, Kopecky demonstrated that the index T_0—the mean lifetime of sunspot groups--which he had introduced changes during the long-term cycle.

Certain authors have assumed that between the 22-year and 80-year cycles there also are cycles with durations of about 33 and about 44 years. The first of these, therefore, consists of three 11-year cycles and the second of four such cycles. Schuster, who used periodograms in an analysis of Wolf numbers, obtained 33.375 years as one of the periods. As usual when periodograms are used, this result to a considerable degree is formal. Clough, in Ref. 62, analyzed solar cycles from 300 A.D. using Fritz's data on auroras (data later serving as a basis for Schove's work). Clough believes that there is some real relationship between the height of the 11-year cycle and its length (Ref. 63). Clough's opinion on this matter was disputed by Eygenson (Ref. 47). In actuality, as already mentioned relative to a different point, the negative correlation between the length of the ascending branch of the 11-year cycle and its intensity at the maximum is compensated by a positive regression of this height and of the length of the descending branch. However, since the second dependence is expressed less strongly than the first, there is a slight negative correlation between the intensity of the 11-year cycle at the maximum and the total length of the cycle.

We have given particular attention to this problem because Clough in his study considered it together with the postulated existence of a 33-year cycle. Using the lengths of past cycles, an attempt was made to determine their height and then use periodograms to obtain long-period components, including a 33-year component. At this point it can be noted that in 1901 Lockyer postulated the existence of a 35-year cycle in the lengths of the ascending branches of the 11-year cycles. However, this hypothesis was not confirmed. It can be said in summarizing that there are no convincing arguments in support of the existence of 33-year (or close to 33-year) cycles. A 44-year cycle was noted first in geophysical phenomena. Köppen discovered it in the alternation of severe winters in Western Europe (Ref. 64) and Cullmer in the frequency of low-pressure paths over the United States (Ref. 65).

An attempt to detect a 44-year cycle directly in solar phenomena was undertaken by A. Ya. Bezrukova (Refs. 66, 67). The annual values of the mean sums of sunspot areas, taken separately for the solar northern and southern hemispheres, yield curves of similar shape for the same hemisphere after two 22-year cycles, i.e., after a mean period of 44 years. If even and odd cycles are combined by the method proposed by Gnevyshev and Ol' and a study is made of the ratio of odd and even 11-year cycles for the year before the maxima, this ratio, as a function of the numerical disignation of the pair of 11-year cycles, reveals a clearly defined four-cycle, that is, 44-year "period". This conclusion is of interest not only as confirmation of the reality of the existence of a 44-year cycle, but also as evidence of some particular property of the year preceding the maximum of the 11-year cycle, in most cases coinciding with the first knee on the Kozik-Kopecky curve.

Since the result obtained in the abovementioned studies was obtained analyzing solar activity separately for the two hemispheres, we will now discuss the problem of change of the ratio of activity of the solar northern and southern hemispheres over an 80-year cycle and change of the heliographic latitudes of zones of spot formation within this cycle, although this problem, to be sure, will be considered at considerable length in one of the later sections of this chapter in connection with the 11-year cycle.

It was demonstrated (Ref. 68) as early as 1941 that during the period from 1865 through 1910, activity predominated in the southern hemisphere, and, in the period from 1910 through 1938, in the northern hemisphere. As a result, it was postulated that there is a 70-year cycle in the activity of a particular solar hemisphere. The annual values of the ratio $\frac{N - S}{N + S}$, where N and S are the mean annual spot areas in the solar northern and southern hemispheres, respectively, reveal an 80-year cycle in the ratio of the activity of the solar hemispheres (Ref. 69).

Symmetry of solar activity, that is, cases in which the level of activity in both hemispheres is more or less identical, occurs when there are extremely high or extremely low 11-year cycles.

Maximum asymmetry is characteristic of 11-year cycles of medium intensity. Ref. 33 contains an interesting diagram which the author (Waldmeier) considers useful as a mnemonic tocl (Figure 11). As shown by the figure, in an epoch of the minimum of the 80-year cycle the activity is distributed more or less uniformly over the solar disk. After 22 years, that is, after completion of one 22-year cycle or a quarter of an 88-year cycle, the activity predominates in the solar northern hemisphere. Twenty-two years later the activity again is distributed uniformly, but its level naturally has increased because the epoch of the maximum of the long-term cycle already has arrived. After another 22 years, that is, 66 years after the commencement of the 80-year cycle, the activity is concentrated in the southern hemisphere.

Finally, after still another 22 years there is a return to the conditions prevailing at the minimum: activity is low and distributed quite uniformly between the northern and southern hemispheres. Smaller circles are shown at the center of the schematic circles in Figure 11; they arbitrarily represent a "nucleus" generating the spots. In an epoch of minimum activity, this nucleus is displaced into the northern hemisphere.

It is postulated that the rate of propagation of a cyclic disturbance from the solar interior to its photosphere is small; therefore, any position of the generating nucleus appears in the photosphere only after passage of some rather considerable time. The predominance of activity in the solar northern hemisphere 22 years after the epoch of the minimum of the 80-year cycle is associated with the presence in the northern hemisphere of the generating nucleus in the epoch of the minimum of this cycle. The predominance of activity in the southern hemisphere 66 years after the epoch of the minimum of the long-term cycle can be attributed to the displacement of the generating nucleus into the southern hemisphere in the epoch of the maximum of this cycle. All these results were obtained by Waldmeier using seven 11-year cycles, Nos. 12-18 inclusive, i.e., on the basis of data not even covering a full 80-year cycle. It is entirely possible, therefore, that further observations will introduce considerable corrections into the initial result. In particular, the development of cycle N. 19, falling in a epoch of the maximum of the current 80-year cycle, shows for the time being that activity is predominating in the northern hemisphere, despite the model shown in Figure 11.

The 80-year solar cycle also is manifested in the heliographic latitudes of the spot groups (Ref. 38). Thus, by using Greenwich data since 1874 and employing somewhat earlier observations by Spörer, the following analysis can be made. If the difference between the mean

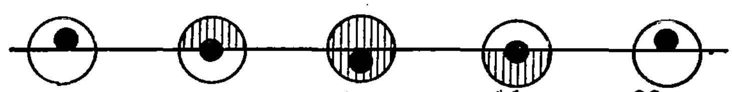


Figure 11. Waldmeier model of distribution of activity in solar northern and southern hemispheres as function of phase of 80- to 90-year cycle

heliographic latitudes in the solar northern and southern hemispheres is determined for each 11-year cycle, and the number of solar rotations during which positive and negative values of this difference were observed is calculated, it is possible to determine the value

$$A = \frac{n - s}{n + s},$$

where n is the number of rotations in a cycle when the difference in heliographic latitudes is positive, i.e., the mean heliographic latitude in the northern hemisphere is higher than in the southern hemisphere; and, s denotes the number of rotations with a negative value of the difference, i.e., it indicates the number of cases of predominance of higher latitudes in the solar southern hemisphere. The results of such calculations have been given in Table 9.

On the basis of this table, the conclusion has been drawn that there is a cyclic variation of the latitudes of spots; the duration of the cycle is close to 80 years. It is obvious, however, that no final conclusions can be drawn on the basis of these data, which cover only a single 80-year cycle. It is true that the Gleissberg criterion (Ref. 70), applied to these data, indicates the reality of the long-term variation reflected in Table 9, but this criterion in itself does not have an adequately sound basis (Ref. 71). The data in Table 9 correspond to the model shown in Figure 11, but this is not unexpected, since both the Table and Figure 11 are based on the same data.

In concluding this section we will note a curious fact in the history of solar observations: between the years 1672 and 1704 no spots were observed in the solar northern hemisphere (Ref. 47). It is of interest to relate such an extreme case of asymmetry, persisting more than a full 11-year cycle, to the corresponding phase of the long-term cycle. By using Table 6, as well as data on the epochs of the first telescopic

Table 9

No. of cycle.	10	11	12	13	14	15	16	17	18
A	0.50	0.11	-0.36	-0.34	-0.23	-0.15	0.10	0.12	0.20

11-year cycles, we conclude that 1672-1704 falls at the very end of the 80-year cycle No. 1 and the beginning of cycle No. 0 (that is, the long-term cycle having a maximum in the 18th century). Therefore, for the greater part of this epoch, specifically for 1672-1698, the conditions were those associated with the end of the descending branch of the long-term cycle, characterized by a predominance of activity in the solar southern hemisphere. This extreme case of asymmetry, therefore, also fits into the abovementioned Waldmeier model, but it must be said that it is impossible to explain such an exceptionally well-expressed asymmetry within the framework of the 80-year cycle. In this case a cycle of still greater duration apparently was manifested.

The next section will discuss the problem of extremely long solar activity cycles.

Section 4. Solar Cycles of Long Duration

There is reason to believe that the "hierarchy" of solar cycles does not end with a cycle with a duration of 80-90 years, but that there are cycles of a greater duration. The study of such cycles obviously is extremely difficult because even for the shortest cycles of this kind the length already is commensurable with the period of time during which solar telescopic observations have been made. The existence of solar cycles whose duration appreciably exceeds the duration of the period of telescopic observations, therefore, can be judged only by use of indirect indicators or on the basis of visual sunspot observations. We will first discuss cycles which can be judged at least in part on the basis of telescopic observations. Oppenheim, in a study already mentioned, feels that solar activity can be represented as a biperiodic function, whose second period is of 180 years duration. As will be shown below, many other investigators have obtained the same period, or one close to it.

P. P. Predtechenskiy has advanced the hypothesis of the existence of a 189-year period of solar activity which he has called the "indiction" (Ref. 50). Although Predtechenskiy's work now is only of historical interest, it is of value to discuss it in somewhat greater detail because some of his ideas have a certain heuristic interest. Predtechenskiy assumes that the nonuniformities in length and amplitude inherent in the 11-year cycle disappear for a cycle with a duration of the order of 190 years, and a 189-year cycle already is a rigorous period.

It is impossible to accept this assertion, since as already pointed out in the preceding section, the 80-90 year cycles also differ in duration and amplitude, and there is no basis to assume that the same will not be true of the 189-year cycle.

A more interesting assertion of Predtechenskiy's is that the 189-year "period" is basically a period of repetition of types of 11-year cycles. According to Predtechenskiy, there are three types of 11-year cycles: (1) symmetric, with rising and descending branches which are close in duration and with a not very great height at the maximum; (2) asymmetric, characterized by an appreciable difference in the lengths of the ascending (shorter) and descending (longer) branches and a great height at the maximum; and, (3) singular, constituting the superposing of an explosion-like peak of relatively short duration on a symmetric cycle.

The indiction consists of seventeen 11-year cycles. Predtechenskiy classifies cycles Nos. 9 and 18 as singular; No. 9 is categorized as a "small" singular cycle; and, No. 18 is called a "large" one. Other than height at the maximum, the difference between them is that the "small" singular cycle is followed by two "asymmetric" cycles in a row, whereas the "large" singular cycle is followed by three consecutive asymmetric cycles. The singular cycles disrupt the alternation of heights in accordance with the low cycle-even, high cycle-odd pattern. Since Predtechenskiy used quartile data he could only employ data since 1749, that is, at the present time only a little more than the postulated 17-cycle "period". There is now no basis for assuming that the Predtechenskiy model is correct, although it does reflect certain properties of the long-term solar cycle. A very weak aspect of his work is the lack of objective criteria, i.e., the assignment of a cycle to a particular category, although several attempts to do so were made.

The existence of singular cycles has been confirmed in a recently published paper by Xanthakis (Ref. 72). This author considers that Nos. 9 and 19 are singular 11-year cycles. In 1954 Anderson (Ref. 73) postulated the existence of a 15-year cycle period of solar activity, that is a "period" of 169 years. He arrived at this conclusion by comparing the variation of mean annual Wolf numbers for 1749-1800 and for 1917-1953. The 169-year period consists of two parts--one of 88 years and the other of 81 years. A study of the old Fritz data confirms, or at least does not contradict, these assumptions. Anderson's work served as a stimulus for a more detailed study of the possibility of the existence of cycles with a duration of several centuries. The possibility of the existence of a cycle consisting of 16 11-year cycles was suggested in Ref. 74. It was also pointed out there that the 169-year cycle can be traced using mean monthly values as well as mean annual values.

Attempts by the author of Ref. 75 to discover a double 80-90 year cycle equal to approximately 15 11-year cycles should be mentioned.

An experiment made by Bonov (Ref. 76) is not without interest; the epochs of the maxima and minima since 1610 were used to obtain the sums of the durations of the ascending branches in even and odd cycles. From the ratios of these sums it was found that there is a 176-year solar cycle consisting of two 88-year cycles, of which the first was the weaker and the second the stronger. Dzhurkovich (Ref. 77) also indicates the existence of a period of 15 cycles. None of these investigations can be considered convincing for a clearly understandable reason. The series of observations at the disposal of these authors did not exceed the duration of the period they sought to discover.

A. I. Ol' (Ref. 78) noted that by connecting on a graph the points of the maxima of the 80-year cycles of the 18th and 19th centuries, and also the points of the minima of the 18th, 19th and 20th centuries, by straight lines, the result is two parallel straight lines having a small inclination to the x-axis and revealing a multisecular increase (of course, within the epoch considered). This is clearly illustrated in Figure 12. It is true that in this case secular smoothing was not done using formula (1.25), but rather by the elementary method which Gleissberg used in his early studies. As shown in the preceding section, this method leads to the conclusion that the long-term cycle of the 19th century is higher than the corresponding cycle in the 18th century. It is impossible, however, to attribute the entire supersecular variation to this effect because the heights of the minima of the 80-year cycles also have been found to have a supersecular variation, and these heights are not dependent on the method of multisecular smoothing.

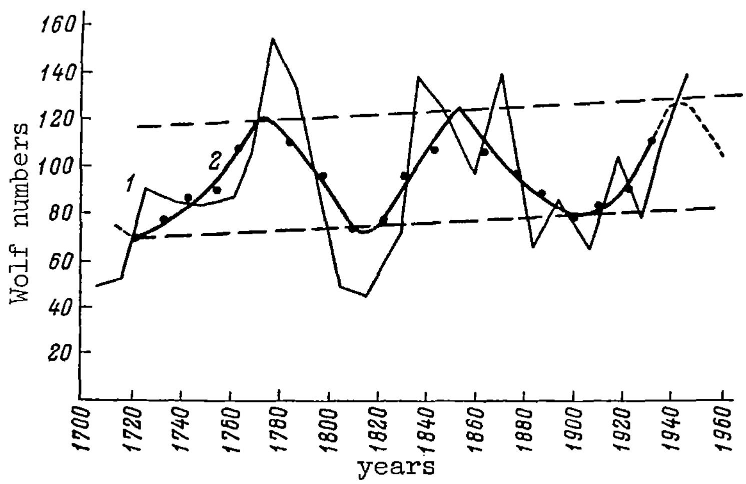


Figure 12. Indication of the existence of solar cycles of a duration greater than 80-90 years: 1, nonleveled points of maximum for 11-year cycles; 2, leveled by the 4 maxima method (according to A. I. Ol')

For cycles of greater duration than the period of telescopic observations, particular mention should be made of Clough's 300-year cycle which had maxima in the 3rd, 6th, 9th, 12th, 13th and 19th centuries. Clough drew this conclusion on the basis of Fritz' data on the frequency of occurrence of auroras. The minimum of the 300-year cycle corresponds to the epoch of maximum ratios of the time interval between successive minima of the 11-year cycles to the time interval between successive maxima. Using this approach it is found that the most recent of the minima of the 300-year cycle was in 1680. This epoch was in fact characterized by low solar activity, but at the same time there is no particular basis for assuming that the 19th century is characterized by particularly high activity; on the contrary, both the 18th and 20th centuries have higher levels of activity. Therefore, we feel that the existence of a 300-year solar activity cycle is dubious. Certain climatologists have cited facts (Ref. 79) supporting the existence of such a variation in solar activity, having found a similar phenomenon in the supersecular variation of hydrological characteristics. However, as will be discussed in detail in Chapter V, multisecular variations of the meteorological and hydrological regimes can be caused by solar activity cycles of still greater duration.

On the basis of investigations (a) of climatic variations, (b) number of discovered comets, (c) auroras, and (d) secular variations of terrestrial magnetism, Link has discovered a 400-year "period" of solar activity. According to Link, the number of discoveries of comets is maximum in the epoch of the minimum of the 400-year cycle, and vice versa. Such a conclusion to a certain degree contradicts the generally accepted point of view that there is an increase in the brightness of comets with an intensification of solar activity (this will be discussed in greater detail in Chapter 6). It is obvious that in epochs of intensified solar activity the nighttime cloudiness associated with cyclonic activity will attain a maximum, thereby decreasing sharply the number of discovered comets, even if it is assumed that their brightness increases with an increase of solar activity. Of course, such a conclusion is correct for the historical past when comets were discovered with the naked eye and especially when astronomical observations were localized to a very high degree. In order to check Link's hypothesis it is possible to use Schove's data; these reveal that the high level of solar activity observed at the present time had an analog in the 16th century, although in certain respects the present state of solar activity is closer to that observed in the 14th century. In the 12th century the activity level was lower than in the 14th and 16th centuries, but higher than in the neighboring 13th and 11th centuries.

Nothing exceptional was observed in the eighth century, but in the fourth century the 80-year cycle was well expressed, which possibly is explained by the general high background of activity and can be related to the 400-year cycle. Thus, there actually are certain hints of the reality of the cycle whose duration was postulated by Link.

In 1949 the author processed the catalog compiled by Denning (Ref. 80), listing the number of comets discovered in each century by the naked eye. Data from the 1st century B.C. to the 19th century were used. The processing involved the application of the simplified Alter periodogram method (Ref. 81) to the numbers of comets included in the catalog. In this way it was possible to establish a tendency to a recurrence of large numbers of comets at periods of 600 and 900 years (Ref. 46). Since the latter "period" falls only twice in the epoch covered by the catalog we are not inclined to attribute great importance to it.

The existence of a 600-year cycle merits somewhat greater consideration because there are additional indications of its reliability. I. V. Maksimov, who investigated 80-year cycles which he discovered in tree rings (as already discussed), found that the amplitudes of the 80-year variations themselves are subordinate to a cycle of still greater duration whose mean duration is 600 years (Ref. 56).

If an attempt is made to discover solar cycles of still greater duration, it inevitably is necessary to deal with data which is neither of solar nor even of astronomical character (such as Dunning's data on comets). Paleoclimatic and paleohydrological data then become of decisive importance, but in order to use them for the purpose of discovering solar cycles of a greater duration it is necessary first to establish the relationship between solar and geophysical phenomena of this type for a shorter period of time. As already mentioned, this is discussed in Chapter 5. At this point we will discuss only the work of A. V. Shnitnikov. Using Shostakovich's data (Ref. 82) on the lengths of 11-year cycles from the 13th century B.C. through the 17th century A.D. Shnitnikov (Ref. 79) supplemented his data for the telescopic period of observations and obtained the mean duration of the cycle.

He then determined the deviations of cycle duration from this norm and constructed an integral curve; he also defined the epochs of maximum and minimum deviations of the lengths of cycles from the norm. During the investigated period he found time intervals of the order of hundreds of years when the 11-year cycles were longer than the norm, and periods when they were shorter than the norm. Positive deviations predominate from the 12th through the 4th-3rd centuries B.C. The pattern then changes sharply, and in the 1st-2nd centuries A.D., there were clearly expressed negative anomalies which in general continued to the 8th century A.D. Positive anomalies then predominate to the 14th-15th centuries, later being replaced by negative anomalies. The periods of short 11-year cycles should reflect epochs of high solar activity, and the periods of long cycles, epochs of low solar activity. In general, Shnitnikov obtains a cycle with a duration of about 1,800 years.

However, as already noted, the relationship between the general duration of the 11-year cycle and its intensity, if there is one, is expressed very weakly. In addition, the durations of 11-year cycles, established in part through manifestations of solar activity in the lower layers of the atmosphere and in the hydrosphere, do not always correctly represent the duration of solar cycles themselves.

At this point we end our review of current concepts concerning solar cycles of great duration. We have not included in this discussion cycles with a length of millions of years whose existence has been postulated by Willett (Ref. 83), and especially by P. P. Predtechenskiy (Ref. 84), or the hypothetical cycle of variations of solar luminosity which was postulated several years ago in a series of articles by Opik. Proceeding on the basis of concepts of the internal structure of stars, he postulated the existence of an extremely long cycle of variations of solar luminosity.

Section 5. Intracyclic Variations and Fluctuations of Solar Activity

The term intracyclic variations is defined as variations of solar activity within the 11-year cycle which have a more or less stable character. This includes variations from one year to five to six years, or somewhat longer. Certain of these variations are so regular that it is possible to speak of cycles of shorter duration than the 11-year cycle.

Solar activity fluctuations is defined as a process which for the most part has an aperiodic character, and which has a mean duration of several months.

The development of an active center represents a process which is local with respect to heliographic coordinates and which for the most part transpires in a particular sequence. The total duration of this process also is several months or several solar rotations, the solar rotation apparently being a unit of time more natural for determining duration of such phenomena.

We will begin with a discussion of intrasecular variations. In the previously mentioned study of Kimura (Ref. 21), it was shown that there are short (less than five years) "periods". In general, since the beginning of this century there has been abundant use of the periodogram and harmonic analysis methods in study of Wolf numbers, and to a certain degree in study of other solar indices. One of the principal objectives of such investigators has been the finding of periods of shorter duration than the 11-year cycle. Communications have been published on the discovery of periods of two years (Schuster, Ref. 85); 6, 8, 17, 23.5 and 37 months (M. and J. Labrouste, Ref. 86), 2.6 years, 5-6 years, 5.7 years, 7 years, etc.

Considerable attention has been drawn to the 5-6 year cycle postulated by many investigators. As pointed out by M. S. Eygenson, this cycle was not discovered initially on the sun. We feel that the 5-6 year cycle manifested in a number of meteorological and other terrestrial phenomena is expressed poorly in solar processes and cannot be considered finally established in solar processes on the basis of those solar activity indices now at the disposal of investigators. The presence of a second peak of the $\bar{a}$ index, noted by L. A. Vitel's (Ref. 87) and confirmed by M. S. Eygenson and I. A. Prokof'yev (Ref. 88), is expressed poorly and by no means appears in all 11-year cycles (it is true that the latter fact can reflect the singular long-term variation of the properties of this index), and exerts no appreciable influence on its cyclic curve. However, many arguments can be advanced to show that the 5-6 year cycle cannot be reduced solely to a property of certain geophysical phenomena and that instead of a simple 11-year solar wave there is a double wave. First, in recent decades there has been observed a tendency to a second peak in the curve of geomagnetic activity in the 11-year cycle (Ref. 89) and second, there is a 5-6 year cycle in the brightness of comets, a phenomenon definitely related to solar activity. Both these examples suggest that the 5-6 year cycle is a wave in the 11-year variation of the intensity of solar corpuscular radiation.

With respect to the 5-6 year cycle in meteorological phenomena it is impossible to omit reference to a series of investigations of this problem made by Baur (Refs. 90, 91). We will give particular attention to these investigations in Chapter 5, but at this point we will note only that as the explanation for such a phenomenon Baur advanced the hypothesis of variations in the transparency of the solar atmosphere during the 11-year cycle, which is not very sound from the astrophysical point of view. It is far more probable that the 5-6 year cycle is caused by a double cyclic (in the 11-year cycle) wave of rate of change of solar activity. The curve of the first differences of Wolf numbers naturally has two extrema within the 11-year cycle. One is positive, corresponding to the maximum rate of increase of activity, and the other is negative, reflecting the maximum rate of decline of the cyclic curve. Certain geophysical phenomena can be concentrated near such extrema of the first differences of relative spot numbers. We repeat, however, that this is correct almost exclusively for meteorological phenomena.

With respect to shorter "periods", such as 2-3 years, for the time being there is no basis for assuming them to be real; instead, these are the results of formal use of various harmonic and periodogram methods. We already have mentioned the well-known misuse of such methods in the study of the 11-year cycle. A recently published study by Rima (Ref. 92) contains considerable data based on periodogram investigations of solar activity. Rima asserts that "periods" shorter than 11.2 years are stable and real. However, this is scarcely true. The results obtained by Rima nevertheless are of a certain interest. Rima assumes

that the "periodicity" of solar activity is of external origin and is associated with the influence of the planets, especially Jupiter. We already have touched on this problem in the discussion of the 11-year cycle and we will return to it again in Chapter 3. It should be noted that, in general, attempts to discover "periods" shorter than the 11-year cycle very often are induced by a desire to detect in solar activity some frequency which could be related to the periods of revolution of one of the planets. This applies to a series of papers by Link and his associates (Ref. 93).

We will now proceed to the problem of solar activity fluctuations. Historically, the first attempt to find any periodicity in intensifications and attenuations of solar activity, usually lasting several months, is the work of Unterweger (Ref. 94). Unterweger used Wolf number data for 1880-1887 to find the "period" of variations of the values of this index--69.4 days. The first thorough investigation of this problem was made by Frenkel' (Ref. 95), who investigated Wolf numbers for the years 1877-1911. Despite the great care with which this work was done the result was ambiguous and Frenkel' could draw only preliminary and hypothetical conclusions.

In all probability the study suggested a "period" of 200 days; such a variation was observed in all of the 11-year cycles considered, but only in those cases when solar activity did not fall below a set level. In some cases this "period was only 150 days. Its amplitude is less by a factor of 10 than the amplitude of the 11-year cycle; the "period" of 68.5 days is very close to the abovementioned "period" obtained by Unterweger. In Ref. 2 Gleissberg expressed doubt as to the reality of the "periods" noted by Frenkel'. Later the problem of solar fluctuations was studied from various points of view. A variety of methods was used in detecting solar fluctuations, such as the following. The difference was obtained between the observed Wolf numbers and Wolf numbers smoothed using formula (11) (Refs. 38, 39). In the curve of these differences it is possible to detect an 80-year cycle, but efforts to discover a shorter period were unsuccessful.

Yu. I. Vitinskiy (Ref. 96) made a special investigation with the objective of checking the reality of short periods in solar fluctuations. He analyzed the mean monthly Wolf numbers from 1844 through 1957. The index he used was the ratio of the observed mean monthly value to the smoothed value taken from a cyclic curve constructed using annual values. Epochs of the minimum were not taken into account because in these epochs the value of the ratio used as an index can be exaggerated very greatly artificially. The autocorrelation method and a simplified periodogram method were used to detect the cycles (Ref. 81). No periodicity in the limits of 1 to 14 months was discovered. There is, therefore, basis for assuming that solar fluctuations are an aperiodic phenomenon.

However, if no definite pattern can be discovered in the alternation of fluctuations, this by no means proves that there is no pattern in the development of each individual fluctuation. Two phenomena must be separated here. The first is the synchronous or asynchronous aspect of intensifications of solar activity in different parts of the solar disk. It sometimes happens that the entire fluctuation is caused by the development of only a single strong active center, but sometimes two or more such centers develop simultaneously. The problem of which of these models applies in the development of a specific fluctuation remains open to the present time. In 1940 M. S. Eygenson and the author demonstrated that the most probable process is the simultaneous development of two active regions, but the statistical soundness of this conclusion is not obvious (Ref. 97). Second, the pattern of development of each individual active center (region, focus) is quite definite; the corresponding model is cited in a book by Kuiper (Ref. 4) and has been elaborated in a recently published article of de Jager (Ref. 98); it, therefore, will not be described in this book.

Also of great interest is the complex of geophysical phenomena associated with the development of an active center on the sun. It is understandable that the geophysical phenomena associated with solar flares (sudden ionospheric disturbances, magnetic storms with a sudden commencement, etc.) should be observed most frequently on the first days of development of an active region or in any case during the first rotation of the sun, i.e., in the same rotation during which the active region developed. On the other hand, such phenomena as magnetic storms with a gradual commencement are associated, as is well known, with M-regions, and should coincide with rather late rotations of the sun, i.e., fall in the rather late phases of development of an active center. It, therefore, is entirely possible that a considerable part of the M-regions constitute late stages in the existence of an active center. It goes without saying that these phenomena do not exhaust the full variety of invisible regions of activity (M-regions), which can be formed immediately in that form, bypassing the stage of visible activity.

It can be considered established that the dimensions and the lifetime of an active center do not change appreciably during the 11-year cycle. With development of the cycle the principal change is in the number of active centers. This in part explains the small cyclic amplitude of such indices as the AGA and the large amplitude of such indices as Kopecky's f_0 index.

Section 6. The Statistical Character of Solar Activity

It was demonstrated in the preceding sections that solar activity is characterized by a large number of cycles of different duration. If we proceed on the basis of what has been established by use of the Wolf

number index, and discard the most doubtful cycles, we still have approximately the following set of cycles whose reality is confirmed formally by means of the following criteria:

1. 17.6 months, established by Labrouste by the periodogram method;

2. about 2 years, established by Schuster by the same method and confirmed by Labrouste as a period of 23.5 months;

3. about 3 years, which Rima cites, apparently having in mind the period of 37 months established by Labrouste using the same method;

4. a period of 4 years, mentioned by Rima, discovered by the periodogram method and expressed quite clearly on the Rima diagram;

5. periods of 4.788 and 8.344 years, established by Schuster by the periodogram method;

6. a series of periods shorter than 5 years, established by Kimura by harmonic analysis methods; among these are periods very close to the abovementioned periods obtained by Labrouste, Schuster and others.

In addition to these "periods", established using Wolf numbers, the existence of the following has been postulated:

7. 3-4 years, on the basis of variations in the number of prominences, noted by Lockyer (Ref. 99);

8. a 5-6 year cycle of the first differences of Wolf numbers.

In addition, we have:

9. the well-known 11-year cycle;

10. the 22-year cycle, established by a number of authors, beginning with Turner and Hale and his associates;

11. 33.375 years, noted by Schuster using the periodogram method, but very poorly expressed;

12. a 35-year "period", indicated by Rima, possibly the same which Schuster has defined with a period of 33.375 years;

13. an 80-90 year cycle, discovered by Wolf and investigated by Gleissberg, Eygenson, Waldmeier and others;

14. a 169-year cycle, discovered by Anderson;

15. a 176-year cycle, discovered by Bonov;

16. a 400-year cycle, discovered by Link;

17. a 600-year cycle, discovered by the author.

We will not discuss possible cycles of greater duration because their detection naturally involves the use of indirect data rather than data from solar observations. The natural problem of the relationship of these cycles to one another has long attracted the attention of investigators. We already have mentioned the work of Oppenheim in which an attempt was made to represent solar activity as a double or multi-period function, but none of these attempts have much validity because, in general, periodogram methods based on a postulated rigorous periodicity do not correspond to what actually occurs in the case of solar activity.

In this connection we already have cited Waldmeier, and we wish to emphasize once again that in this respect he is entirely right. But if there is no "hierarchy of periods", there possibly is a "hierarchy of cycles". M. S. Eygenson answers this problem affirmatively. He assumes that at the basis of solar cyclic behavior, there is an extremely low-frequency variation, i.e., a variation having an enormous period.

The entire observed variety of cycles, or "rhythms", as Eygenson calls them (Ref. 75), represents the manifestation of slow deep-seated processes. The following arguments are advanced for proof of the applicability of such a hypothesis. Any cycle of a greater duration is, in fact, the sum of cyclic processes of a shorter duration. To be sure, the word "sum" is not used in the literal sense. For example, the 11-year cycle is made up of the development of individual active centers, the 80-year cycle of 11-year cycles, etc. Furthermore, Eygenson assumes that there is a morphological similarity of cycles of greater duration, on the one hand, and of shorter duration, on the other. Third, the mean duration of each successive "rhythm" in the sequence is two or three times the duration of the preceding "rhythm" in the sequence.

We will now consider these concepts. We will begin with a determination of the character of solar activity as a multirhythm phenomenon. What is the significance of such a determination?

An investigation of the temporal series of solar, geophysical and other indices leads to the following natural classification of the processes which they express (Ref. 100): (1) random variations; (2) disturbances; (3) rhythmic variations; (4) periodic variations; and (5) latent periodicity.

We will now describe these processes briefly.

Random variations are independent relative to time. By investigating a time series using a correlogram, it can be found that random variations virtually drop out and their effect is manifested only in a dispersion whose value is of the order $1/n$, where n is the number of terms in the series.

Disturbances also are of a random character, but within some interval of time they reveal a relationship to one another. With correlation after an interval of time t, there is noted a gradual attenuation of the relationship, making it possible to establish the spectrum of dimensions of the disturbance.

Rhythmic variations are an alternation of disturbances of different signs. There is a certain mean period, but each successive disturbance is determined only by the end of the preceding one; after the passage of a certain time there can be accumulated any phase difference in comparison with the variations having a period equal to the mean.

Periodic variations can be considered as rhythmic, with rigorous alternation of identical cycles.

Latent periodicity is represented by individual cycles which differ in length and intensity as a result of superposing of variations, disturbances or rhythms on periodic terms. This type of temporal processes sometimes is called "a period with disturbances". Although such a definition is admissible, it somewhat detracts from the concept of latent periodicity because the true period can be distorted not only by disturbances, but also by random variations and rhythms.

Each of these types of processes is reflected differently on a correlogram. The behavior of random variations already has been discussed. Disturbances are represented by curves of the exponential type with a definite decrement of attenuation. Rhythms give waves whose period is equal to the mean period of a variation and the maximum ordinate occurs when $t = 0$; the amplitude gradually decays after intervals of time equal to the mean period and the phases will be different. Periodic variations give regular waves on a correlogram; their amplitude is identical each time when $t = 0$, and in this case the phase also is identical. Latent periodicity at first glance gives a correlogram pattern similar to that which represents rhythms, but, in actuality, the difference between them is appreciable. For an explanation we turn to Newcomb's discussion, but we now note that according to the generally accepted opinion, the 11-year solar cycle belongs to a period with perturbations. The concepts of Newcomb (Ref. 101) leave no doubt concerning this matter.

Newcomb distinguished two classes of periodic phenomena (we note that neither of these cases fall in a category of rigorous periodicity). In the first case there is a principal period on which are superposed

certain irregular effects. Then the onset of each particular phase can be accelerated or delayed; irregular effects do not influence the primary factor causing the principal periodicity. In this case the same phase will return after approximately equal long epochs, but only approximately, and the deviations of the epoch of return of a particular phase from that which would correspond to equal intervals will have the character of random errors. If the length of the true period is P, after the nth period the observed phase will differ from nP by the value $\pm\epsilon$. The value ϵ will not be dependent on n. In the second case, when there is a mean period, this period itself is subject to changes, that is, if the onset of any phase is accelerated at any time this acceleration will be passed on to all succeeding phases. If the probable acceleration or lag in the onset of a particular phase is denoted $\pm\epsilon$, after the passage of n periods this deviation will be $\pm\epsilon\sqrt{n}$. The entire series of years for which there are data on the epochs of the extrema of the 11-year cycles was broken down into three groups: (1) 1610-1720; (2) 1720-1820; and, (3) 1820-1900 (Newcomb's work was published in 1901). By determining the mean duration of the cycle at 11.13 years Newcomb found the computed times of onset of the maxima and minima and computed the deviations from them of the true times of onset of these phases. In the mean deviations there was found to be a definite tendency which Newcomb assumed was related to the low reliability of the observations for 1755-1790. After excluding this epoch from consideration, Newcomb found the following mean deviations: for epochs of maxima the observed epoch minus the computed epoch = +0.3 year; for epochs of minima this same difference was -0.1 year. On introducing these corrections into the observational data, Newcomb obtained small values of deviations without any systematic character. Therefore, it was concluded that the first hypothesis which corresponds to a period of 11.13 years with perturbations, is correct; whereas the second model corresponds to rhythms. A. I. Ol' recently repeated Newcomb's investigation using new data. Data for 1700-1957 were used in the study. He used the same mean period as Newcomb did, i.e., 11.13 years.

The deviations of the observed epochs of maxima and minima of the 11-year cycles from the corresponding computed epochs were averaged over four cycles. No systematic increase of the values of the deviations with an increase of the number of the cycle was observed, which again confirms the correctness of a model of a "period with perturbations", not a model of "rhythms". It was found that there is a regular alternation of positive and negative values of the deviations of the observed epochs of maxima from the computed epochs of maxima when there is averaging over 4 cycles. This and other facts led A. I. Ol' to the hypothesis that the deviations are not entirely random. They are determined to a certain degree by the general level of solar activity, whose change reveals a cycle of a longer duration than 11 years. In this sense it actually is possible to speak of the effect of cycles of a

longer duration, such as 80 years, on 11-year cycles; but this effect does not lead to a distortion of the 11-year "period" itself, that is, its transformation to a rhythm, but to the appearance of those indicators which we have related to perturbations. If the data in Table 6 are processed by the same method which was used by Newcomb and Ol' to process data on the 11-year cycles, it can be concluded that the 80-year cycles are in fact cycles and not rhythms.

In actuality, assuming with Gleissberg (since we used his table) that the length of the "long-term" cycle is 7.1 times the length of the 11-year cycle, and applying the Newcomb method, it can be shown that the differences between the real and the precomputed commencements of the 80-year cycles do not reveal any systematic increase, and in this respect behave much like the corresponding differences of the epochs of the 11-year cycles. Therefore, it can be concluded that the long-term cycle also is a period with perturbations. The small amount of data given by Gleissberg unfortunately does not make it possible to repeat the investigation in the same way as was done by Ol'.

An investigation of fluctuations of solar activity made by Yu. I. Vitinskiy revealed that, statistically, these fluctuations should be assigned to the category of random variations. The fluctuations of solar activity do not reveal definite patterns, either with respect to their alternation or their relationship to one another, or even internally. For example, it is impossible to establish any relationship between the time of increase of a fluctuation and its height at the maximum. Somewhat more definite properties were revealed by repeating spot groups investigated by Vitinskiy in Ref. 96. In that case there was a positive correlation between the time of increase and maximum area. Certain properties of the phenomena and cycles of different magnitudes have been shown in Table 10.

As shown by this table, the 11-year cycle in its internal characteristics differs from the processes of both greater and lesser scale.

The following can be said in summary.

1. There is no such thing as solar phenomena with a "multirhythmic" character. Among these phenomena of different scales, which today can be assigned to a definite category of statistical processes, there is not one which falls in the "rhythm" category. There is such a thing as a "multicyclic" solar phenomenon, but even in this case there is a need for caution. In this respect the only thing which can be considered established is the discovery by A. I. Ol' of a dependence of the epochs of onset of particular phases of 11-year cycles on the cyclic character of a process of somewhat greater duration. This conclusion is confirmed indirectly by the fact that both the 11-year and 80-year cycles fall in the same category of processes, i.e., are periods with perturbations.

Table 10

Type of process	Scale of process	$r_{t, M}$	$r_{\tau, M}$	Index used
—	repeating group	+ 0.70	+ 0.35	area at maximum
random variations	fluctuation	+ 0.33	+ 0.33	amplitude of fluctuation index
period with perturbations	11-year cycle	- 0.81	+ 0.43	Wolf numbers
period with perturbations	80-year cycle	+ 0.58	+ 0.27	Wolf numbers

2. Table 10 shows that great caution must be exercised in approaching the problem of the morphological similarity of cycles of longer and shorter duration (scale). For example, statistically the 11-year cycle cannot be a process of a higher scale relative to fluctuations of solar activity because the latter do not fall in the category of cyclic phenomena, and must be assigned to random variations. Furthermore, the internal relationships of the 11-year and 80-year cycles are completely different. Whereas the first is characterized by a clearly expressed negative correlation between the time of increase and height at the maximum, the second is characterized by a relatively poorly expressed positive relationship between these parameters.

The cited internal characteristics of the 11-year cycle also cause it to differ appreciably from fluctuations, which, in general, reveal no relationship between intensity and time of increase or decline, and from repeating groups, where there is a sufficiently clearly expressed positive correlation between the time of increase and maximum development.

3. The relation between the mean durations of cycles of different duration reveals a tendency to doubling of the length of the cycle with transition to the cycle of next greater duration (scale). However, this relation, insofar as can be judged from available data, does not apply beyond the double 80-year cycle, close to the Anderson-Bonov cycle.

Section 7. Distribution and Cyclic Change of Solar Activity With Heliographic Latitude. Natural Meridional Movements of Active Solar Formations

Soon after the discovery of sunspots it was established that they are observed only in two latitude zones, one in each hemisphere, extending from heliographic latitude ±5 to ±35-40° (Ref. 5). The latitude zone directly adjacent to the equator is avoided by spots and spots appear there very infrequently. Similarly, spots rarely are observed at latitudes greater than 45°. The most high-latitude groups have been observed at latitudes close to 70°. In recent decades high-latitude groups have been observed more frequently which are associated with more powerful 11-year cycles. This fact gave U. Becker and Kopecky basis for speaking of the existence of a second, high-latitude zone of sunspots (Refs. 102, 103). However, Kopecky later reversed himself (Ref. 104) and refuted the results obtained by U. Becker.

Figure 13 represents the distribution by heliographic latitude in the solar northern and southern hemispheres of spot groups with an area of >500 millionths of a hemisphere for groups observed during the period from 1874 to 1953. In the northern hemisphere the maximum number of such groups was observed at latitudes from 0 to 15°; in the southern hemisphere they are encountered for the most part in the zone from 10 to 20°. The distribution of 55 still more significant groups, with an area of $>1{,}500$ millionths of a hemisphere, observed during this same time, is given in Table 11.

As shown by the table, these groups also were concentrated for the most part in the 10-15° zones of both hemispheres. The number of very large groups in the northern and southern hemispheres is virtually the same.

The zone in which solar faculas appear is 15° wider than the main zone of spot formation (Ref. 4). The broadening of the facula zone in comparison with the spot zone is for the most part in the direction of the poles of the corresponding hemispheres. As a result, the mean latitude of faculas is somewhat greater than for spots. High-latitude faculas also occur. They are observed at heliographic latitudes greater than 40°, sometimes as far as 70-75°. They manifest virtually no traces of an 11-year cycle. It was Mascari (Ref. 105) who was the first to mention the existence of these formations, sometimes called polar faculas. Since 1951 polar faculas have been observed rather frequently, especially near the epoch of minimum solar activity in 1954. It should be noted that the precise determination of the latitude of polar faculas is difficult because of the perspective; a very strong influence is exerted in this respect by the inclination of the solar equator to the ecliptic. The most probable mean latitude of the zone is 67°. According to Waldmeier, high-latitude faculas constitute

Table 11

from -30 to -25.1°	from -25 to -20.1°	from -20 to -15.1°	from -15 to -10.1°	from -10 to - 5.1°	from - 5 to - 0.1°	from 0 to 4.9°	from 5 to 9.9°	from 10 to 14.9°	from 15 to 19.9°	from 20 to 24.9°	from 25 to 29.9°	from 30 to 34.9°
3	4	4	10	5	1	0	3	10	6	5	1	1

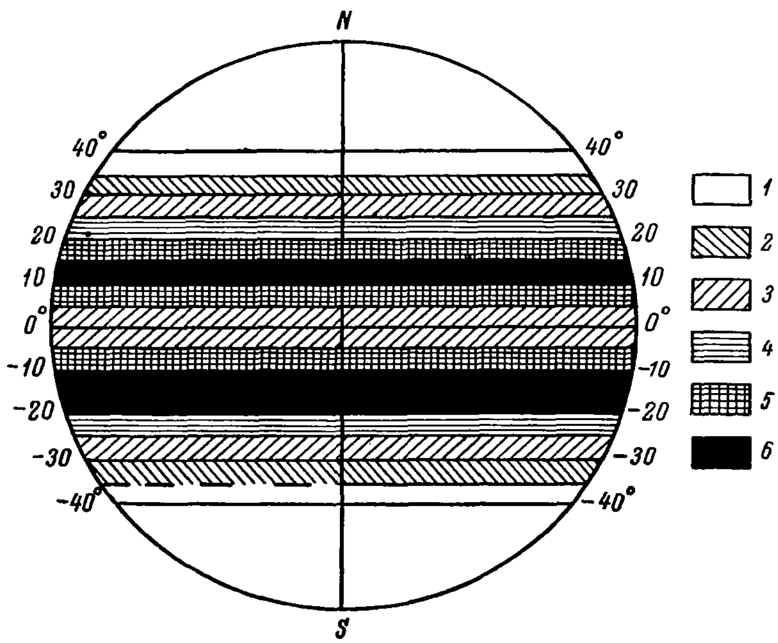


Figure 13. Distribution by heliographic latitude of spot groups with an area of > 500 millionths of a hemisphere during the period 1874-1953. Number of groups: 1 — 0; 2 — 1-9; 3 — 10-19; 4 — 20-49; 5 — 50-99; 6 — 100-149

disseminations of small facula formations, each with dimensions of about 2,000 km (about 3"). The appearance of a considerable number of polar faculas near the epoch of the minimum of cycle No. 19, the most intense 11-year cycle in the period of telescopic observations, and the appearance in this same cycle of the maximum number of high-latitude spot groups (43 of a total of 80 in almost 7 cycles) provides

a basis for assuming that the intense development of high-latitude activity constitutes a characteristic of intense 11-year cycles lying at the "crest" of long-term cycles. Unfortunately, because of the lack of observations, this assumption cannot be checked properly using data for the 19th century.

A relationship has been discovered between the form of the solar corona characteristic of an epoch of the minimum in the cycle and high-latitude faculas. Prominences have a singular distribution by heliographic latitude. Two zones are clearly manifested: a low-latitude zone, displaced like the main facula zone by approximately 10° relative to the main spot zone in the direction of the pole of the corresponding solar hemisphere, and a high-latitude zone, appearing at particular phases of the 11-year cycle at a latitude of approximately 50°; however, the mean heliographic latitude of this zone is still greater.

The position of zones of prominences and their displacement with a change in the phase of the 11-year cycle is represented clearly in Table 12, which was taken from a study by Waldmeier (Ref. 5). The table, for the years 1912-1929, gives the distribution of the area of prominences in 10° intervals of heliographic latitude. At the beginning of an 11-year cycle the heliographic latitude of the center of gravity of the zone of prominences is 40-50°. With the development of the cycle there is a splitting into two zones, one of which is displaced toward the high latitudes and the other toward the low latitudes. The first reaches the polar regions approximately in the epoch of the maximum of the 11-year cycle, and the second, with some lag relative to this epoch, is displaced into the low latitudes. However, on the basis of data for almost two 11-year cycles, there has been no observed case in which the center of gravity of this zone dropped below latitude 10°. If the areas of prominences for a year for each of the lines in Table 12 are simply added, no separation into two zones is observed, as can be seen from the last column of this table. In this case we find only a single zone at latitude 20-40°, which is approximately 10-15° greater than the latitude of spot formation.

The latitudinal distribution of filaments naturally is similar to the distribution of prominences (Ref. 106). The distribution of particularly eruptive prominences has been shown in Figure 14 (Ref. 107). The index used here was the number of prominences in a 10° position angle interval. The scale for the conversion of the distance of the histogram points from the solar limb to the number of prominences in such an interval is given in the figure itself. It can be seen that the existence of a high-latitude zone in the case of eruptive prominences also can be traced using total data, but it would be premature to assume that for eruptive prominences the high-latitude zone, at least in the northern hemisphere, is expressed more clearly than for all prominences (quiescent + eruptive, with the first considerably greater than

Table 12

φ	Year																		
	1912	1913	1914	1915	1916	1917	1918	1919	1920	1921	1922	1923	1924	1925	1926	1927	1928	1929	
80-90°	1	1	1	2	4	27	61	1	2	1	1	2	1	1	27	51	5	2	191
70-80	1	2	3	4	24	138	47	3	2	2	1	1	4	3	162	92	15	5	509
60-70	2	1	6	30	187	105	7	1	4	31	1	6	39	96	182	53	68	38	857
50-60	25	24	93	238	112	72	51	22	13	97	69	130	172	154	128	106	74	64	1644
40-50	49	51	157	107	81	144	108	67	104	124	162	234	140	102	121	118	73	107	2049
30-40	10	25	64	172	163	231	187	142	138	98	92	33	96	120	225	138	161	145	2240
20-30	11	12	31	140	114	223	198	260	155	108	50	34	45	100	225	126	188	227	2247
10-20	16	6	16	81	99	172	181	228	192	108	52	27	31	51	130	117	181	156	1844
0-10	6	3	11	23	58	173	178	167	121	70	44	14	6	23	114	85	90	128	1314

Note. All values ϵ are in prominence units.

the second). The indices used in Table 12 and in Figure 14 are different, and in addition, in Table 12 the results have been summed for both hemispheres.

We will now proceed directly to the problem of changes in the heliographic latitude of active solar formations in the 11-year cycle. We have already touched upon this problem to some degree with respect to the second zone of prominences. Since these changes are investigated most easily in the case of sunspots, and since the corresponding patterns were established first for these formations, our discussion will begin with them.

It was Carrington (Ref. 108) who first called attention to the change, or to be more exact, the decrease of the heliographic latitude of the zone of spot formation. We emphasize that it is the zones of development with which we are concerned; this indicates a change in the position of the zone of spot formation itself with the phase of the 11-year solar cycle. Carrington traced this phenomenon for 60 solar rotations, beginning in 1854, preceding the minimum of the coming cycle of solar activity, and discovered a gradual decrease in the latitude of spot formation with time. A thorough study of this problem was made by Spörer, who in Ref. 109 revealed the universal character of such a displacement. Spörer used old observations up to the time of Scheiner's work, that is, beginning in 1621. The law established by Spörer can be formulated in the following way:

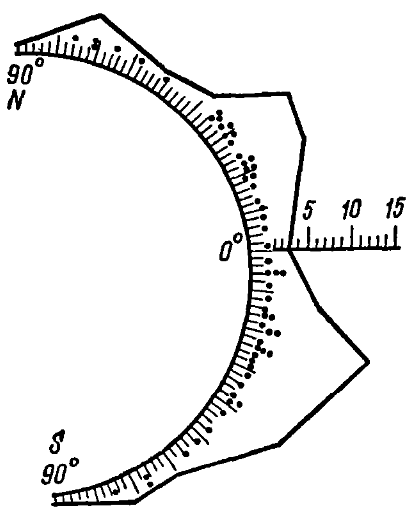


Figure 14. Distribution of eruptive prominences by heliographic latitude (according to Pettit)

1. In an era of a minimum, when the zone of spot formation of the old 11-year cycle attains its lowest heliographic latitude, the zone of spot formation of a new 11-year cycle develops in the high latitudes, gradually descending into the low latitudes, which it reaches near the epoch of the minimum of the next 11-year cycle.

2. The epoch of the appearance of the zone of spot formation of the new 11-year cycle in the high heliographic latitudes and the epoch of the minimum of this cycle according to Wolf numbers can diverge somewhat in time, but not by more than 1 to 2 years.

3. Whereas the curve of change of Wolf numbers with the phase of the 11-year cycle is continuous, the curve of change of the heliographic latitudes of spot groups according to the same argument systematically experiences a break near the epochs of the minima of the 11-year cycles.

All these properties are well illustrated by the familiar "butterfly" diagrams constructed by Maunder (Ref. 110), which have been reproduced in many books. It was Spörer's opinion, also shared by Maunder, that there is displacement of one zone of spot formation in each solar hemisphere. Lockyer, who considered the areas of spot groups in 3° latitude intervals (Ref. 111), held a different view. In such a study are found several "centers of action" or "centers of development" (Ref. 112) of spots which were displaced in the direction of the equator. The method used by Lockyer was criticized by Maunder, who noted that the extent of spot groups in a meridional direction in many cases considerably exceeds 3° (Ref. 113). The work of Lockyer would be forgotten if B. Bell (Ref. 112) had not returned recently to similar concepts. Bell studied the magnetic fields of spot groups and concluded that Lockyer was correct, and that, in actuality, there is a series of spot-forming zones on the sun having the form of narrow bands ("caterpillars", to use Bell's terminology). This problem was investigated recently by Yu. I. Vitinskiy. He demonstrated the untenability of Bell's contention and his results (Ref. 114).

Spörer, as already mentioned, traced the heliographic latitudes of spots in the early telescopic epoch, although on the basis of extremely fragmentary data, and discovered that at the transition from the 17th to the 18th centuries, when the level of solar activity was very low, the maximum heliographic latitude of a group was only 19° (Ref. 115).

By a study of seven 11-year cycles, from No. 10 through No. 16, Gleissberg (Ref. 116) established a "period" of change of the latitudes of the spot-forming zone:

$$P_1 = 11.1 \pm 0.4 \text{ years.}$$

During this same time the 11-year cycle has a mean duration

$$P_2 = 11.4 \pm 1.4 \text{ years.}$$

A comparison of these values shows: (a) a noncoincidence of the minima according to Wolf numbers and according to the times of appearance of the spots of the new cycle; and (b) that the mean error P_1 is less than one-third of the mean error P_2. Thus, the Spörer law seemingly describes a more regular phenomenon than the Schwabe-Wolf law; the variations from cycle to cycle in the first case are smaller than in the second case.

It is of interest to study the character of displacement of the spot-forming zones in the 11-year cycle in relation to their intensity. It was established in a study by M. N. and R. S. Gnevyshev (Ref. 117) that on the descending branches of the 11-year curves of Wolf numbers, the value R for a particular year is determined by the value of the heliographic latitude, and when $\varphi \leqslant 12^{\circ}$ there is no relationship between latitude and the intensity of the cycle, i.e., the same Wolf number values always correspond to these latitudes. Table 13 gives these values.

One of the most important conclusions drawn in the study cited is the establishment of a dependence between the intensity of the 11-year cycle and the heliographic latitude of the first spot groups of this cycle. The higher the latitude at which a particular cycle begins, the greater will be the value R_M. Further investigations of this problem revealed that when R_M is plotted on a graph as a function of the heliographic latitude of the beginning of the cycle, we obtain two straight regression lines, one of which corresponds to even, and the other to odd 11-year cycles (Ref. 118, Figure 15). An exception to the rule is cycle

Table 13

8°	10°	12°
11	33	63

No. 13, which in this respect falls in the "even" group. At the same time, as shown by Figure 16, the latitudes at which the 11-year cycles end do not reveal any relationship to cyclic intensity. Thus, an arc passed through the spot-forming zone for a cycle is dependent on the intensity of this cycle.

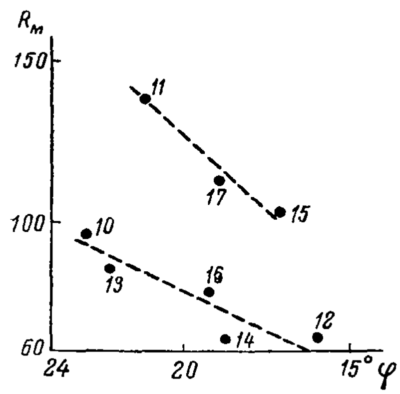


Figure 15. Dependence of the height of an 11-year cycle at the maximum on the heliographic latitude at which the cycle begins. The figures indicate the number of the cycle

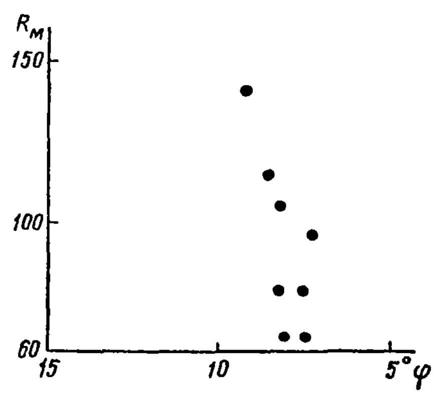


Figure 16. Dependence of the height of the 11-year cycle at the maximum on the heliographic latitude at which the cycle ends

As pointed out by Gleissberg (Ref. 119) and the author (Ref. 120), the rate of displacement of the spot-forming zone from high heliographic latitudes to low latitudes is related to the law of solar rotation. We shall return to this problem in the next section.

We shall now proceed to the problem of the natural movements of sunspot groups in a meridional direction, or as is said frequently, but not altogether accurately, "in latitude".

Carrington (Ref. 121) was the first to mention the presence of definite patterns in this phenomenon. He found that spots situated in a zone bounded by heliographic latitudes ± 15 have natural movements along a meridian toward the equator, whereas spots situated outside this zone are characterized by movements in the direction of the pole of the solar hemisphere where these spots are situated. This problem was studied later by Spörer (Refs. 122, 115). On the basis of study of rather long-lived spot groups, Spörer, in general, came to the same conclusions as did Carrington.

At a later date a thorough investigation of the natural movements of spot groups along a meridian was made by Tuominen (Ref. 123). This author had available far more data than did Carrington and Spörer, specifically, data on long-lived groups according to Greenwich records for 1874-1935. In almost complete agreement with Carrington and Spörer, Tuominen found that groups whose latitudes fall in the range $\pm 16^{\circ}$ have a tendency to be displaced toward the equator and those whose latitudes lie outside this zone have natural movements in the direction of the pole of the particular hemisphere. The scattering of the determined values was large. The rate of natural movements is dependent on heliographic latitude; at latitudes of $\pm 16^{\circ}$ and near the equator, the rate of natural movements becomes zero. The maximum values in the direction of the equator are attained at latitude $\pm 3^{\circ}$. In addition, relatively high rates in the direction of the corresponding pole occur in the highest latitudes at which spots are observed, i.e., about 30°.

The results obtained by Tuominen are shown in Figure 17. Latitudes have been plotted along the x-axis (south latitude at the left, north latitude at right). The rate of displacement of spot groups along the meridian in degrees per day has been plotted along the y-axis; the positive rates in the northern hemisphere denote movement in the direction of the pole and negative rates denote movement in the direction of the equator, while in the southern hemisphere positive rates denote movement toward the equator and negative rates denote movement toward the pole. The solid curve represents observed values; the dashed curve represents smoothed values. A study for this same epoch of the maxima and minima of 11-year cycles separately, and the even and odd cycles separately (Ref. 214) revealed the following: (a) at latitudes exceeding $\pm 7^{\circ}$, movement along a meridian has a tendency to be directed in

opposite directions in successive cycles; (b) at latitudes $\pm 20^\circ$ and 0°, there are no movements along the meridian; and, (c) within the equatorial zone, i.e., within a zone of $\pm 7^\circ$, spots move toward the equator with the greatest velocity at latitude 3°. The conclusions drawn in the preliminary investigation, therefore, are confirmed. These results have been shown in Figure 18. The difference in the behavior of the natural movements of spot groups along the meridian in even and odd cycles led to speculation on the part played by the 22-year cycle in the overall behavior of these groups.

A special study of this problem was made by M. Schwarzschild and Richardson (Ref. 125). On the basis of data for 1,800 repeating sunspot groups, they demonstrated that there is an appreciable variation having a 22-year "period"; in the odd cycles the natural movements are directed toward the equator and in even cycles toward the pole. The half-amplitude of the variation is 0.0043°/day or 60 cm/sec, with a probable error of 0.0009°/day. The problem later was investigated by U. Becker, who established that natural movements along the meridian not only are a function of heliographic latitude, but also are dependent on the position of the spot-forming zone.

The boundary between polar and equatorial directions of natural movements, therefore, should not be related to a definite heliographic latitude; this boundary is unfixed and is itself a zone of spot formation which moves during the cycle. At the same time, no confirmation has been obtained of the postulated relationship between natural movements along the meridian and the phase of the 22-year cycle, but this relationship with the phase of the 11-year cycle is very definite, passing through the zone of spot formation (Ref. 126). These results have been confirmed in a recent study by Tuominen, who used data on the meridional displacement of groups with a short lifetime, separately for the epochs of the maxima and minima of the 11-year cycles (Ref. 127). A critical review of the studies made earlier in this field was made by Yu. I. Vitinskiy (Ref. 128), who investigated groups in the longitude zone from 0 to $\pm 70^\circ$ from the central meridian. The author used 5° latitude zones. Movement in the direction of the equator is observed for such groups at all latitudes. In the limb zone, that is, from ± 70 to $\pm 90^\circ$, there is movement towards the equator at certain latitudes and movement towards the pole at others. It should be emphasized that all this work was done only for long-lived spot groups; for these groups Vitinskiy feels that the movement toward the pole that was observed in the limb zone is an "apparent phenomenon", associated with the inclination of the axes of bipolar groups to the solar equator, with this inclination increasing with an increase of heliographic latitude. The investigation of individual spots in this same limb zone in all cases revealed movement toward the equator. By selecting spots which appeared and disappeared at a distance not greater than ± 70 from the central meridian, Vitinskiy found that for such groups, provided they lie

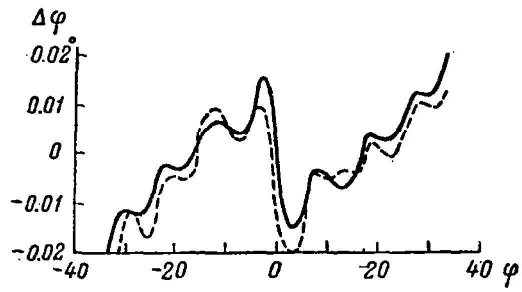


Figure 17. Natural movements of spot groups along the meridian (according to Tuominen)

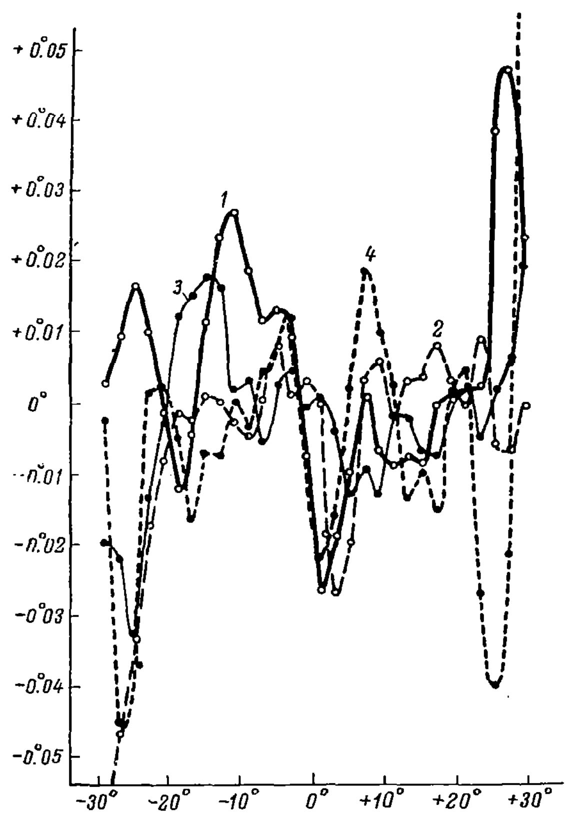


Figure 18. Natural movements of spot groups along the meridian, shown separately for epochs of maxima and minima, and for even and odd cycles (according to Tuominen): 1, maxima of odd cycles; 2, maxima of even cycles; 3, minima of odd cycles; 4, minima of even 11-year cycles

at low heliographic latitudes, the predominant direction of natural movement is toward the equator, but at all other latitudes the predominant direction is toward the pole. The mean diurnal displacement for the northern hemisphere is 0.0175°/day, and for the southern hemisphere, 0.0039°/day. Most data used were for the 20th century.

In summarizing the results of investigations of natural movements of spot groups along the meridian it is possible to note the following:

1. The meridional natural movements of spot groups occur at a rate of the same order of magnitude as the rate of displacement of the spot-forming zone itself (about 1 m/sec); this rate is greater in the high latitudes and minimum in the low latitudes. In this respect there also is an analogy with the rate of movement of the spot-forming zone.

2. The predominant direction of movement of sunspots along the meridian is toward the pole of the hemisphere where these spots are situated. In the low latitudes movement of the spots toward the equator is predominant.

3. The precise boundary between the "streams" directed toward the pole and the equator, respectively, cannot be determined at the present time, but it can be assumed that it lies between latitudes 10 to 15° of the particular solar hemisphere, and probably is related to the position of the spot-forming zone, i.e., to the phase of the 11-year cycle. The postulated relationship with the 22-year cycle still requires further investigation.

4. The results of statistical investigations of this phenomenon can be distorted by various effects, some of which are mentioned in the study of Yu. I. Vitinskiy.

We shall now consider the problem of the natural movements of filaments along the meridian. As we have seen filaments develop for the most part near clusters of large spot groups, but attain their maximum later than the maximum development of spot formation in a particular area of the sun. The lag in the maximum development of filaments relative to the maximum development of groups in a particular active region, as pointed out in Section 5, averages 80 days; the higher the heliographic latitude, the greater is this lag. The displacement of filaments during the time of a solar rotation in the direction of the high latitudes, characteristic of the late stages of development of active centers, is dependent on heliographic latitude. This dependence is represented in Table 14.

Filaments which do not develop in relation to any spot group also are displaced toward the pole of the corresponding hemisphere and at approximately the same rate as shown in Table 14. The movement of

Table 14

Heliographic latitude	Displacement of filaments along meridian during one solar rotation
0-10°	2.3°
10-20	1.7
20-30	1.3
30-40	1.2
40	0.8

filaments usually continues until they reach a heliographic latitude of ±45° (the latitude of the second zone of prominences).

We shall now discuss the "asymmetry" of the solar hemispheres, i.e., the nonuniform distribution of activity in the solar northern and southern hemispheres. This was first noted by Maunder in 1904 (Ref. 129). It was shown that in certain years and even during the course of entire 11-year cycles, activity predominates in one hemisphere and then passes into the other. Greenwich observation data for sunspot areas covered only a period of approximately 30 years, but it was noted that despite a general parallel variation in spot areas, their maxima (cyclic) in the southern hemisphere had a single peak, and those in the northern hemisphere a double peak. The single-peak maximum in the southern hemisphere falls in an intermediate epoch between the primary and secondary maxima in the northern hemisphere. Maunder returned to this problem in 1922 when he could use Greenwich data for 1874-1918, i.e., for the end of cycles Nos. 11, 12, 13 and 14, and a considerable part of No. 15.

It was found that in comparison with the relationship between single- and double-peak maxima in the different hemispheres observed in the earlier epochs, the pattern changed in cycle No. 14. In this cycle two peaks were observed in the southern hemisphere. At a later date asymmetry of the hemispheres was investigated by Brunner-Hägger and Liepert (Ref. 68); they established that the maximum spot area values, considered by hemispheres, reveal a negative correlation in the northern and southern hemispheres. This becomes understandable if we recall the "depression" noted by Maunder, that is, if we take into account that the maximum value of the area in that hemisphere in which a particular 11-year cycle has a single-peak development often falls in the interval between the two maxima of the hemisphere in which the cycle has a double peak. The maximum discrepancy in the epochs of spot area maxima by hemispheres is 1.7 years. Brunner-Hägger later found that

according to Wolf numbers the maximum of cycle No. 17 was double-peaked in the southern hemisphere and single-peaked in the northern hemisphere (Ref. 69). As indicated in Section 3 of this chapter, the asymmetry of the hemispheres is related closely to the 80-year solar cycle. A number of studies of the asymmetric distribution of activity in the solar northern and southern hemispheres was made by A. Ya. Bezrukova (Refs. 130, 131). The principal pattern found is that if the 11-year cycle in one hemisphere has a single-peak character, in the other it will be double-peaked, or to employ the term used initially by Bezrukova, deformed. As a matter of fact, this term is closer to reality, as the second peak sometimes is expressed extremely poorly. It is this property which justifies the application of the term "asymmetry of hemispheres" to the phenomenon. Attempts to establish a regular alternation of normal-single-peaked and deformed-double-peaked cycles in a particular hemisphere have not yet yielded clear results. We feel that Waldmeier is closest to the truth with his rule of alternation of activity (see Figure 11).

Also of interest is Bezrukova's conclusion concerning the existence of a 44-year cycle, based on the asymmetry of the hemispheres. In general, it can be stated that the problem of the asymmetry of the solar hemispheres is still far from solved. An important consideration here is that reliable data on the coordinates of sunspots are available only for the last eight cycles, i.e., these data cover only one long-term cycle. It, therefore, would be somewhat premature to insist on the series of conclusions drawn by the authors mentioned here and by other investigators (Ref. 132). If this is true for sunspots, it is true to a still greater degree for other solar indices. In conclusion, although the existence of a certain asymmetry in the development of solar activity separately for the northern and southern hemispheres is indisputable, it should be remembered that, in general, activity in the northern and southern hemispheres is rather closely correlated. Asymmetry, as already mentioned, is characteristic primarily for epochs close to the maximum of solar activity. For the most part, however, the annual values of spot group areas in the northern and southern hemispheres are correlated rather closely (Ref. 47). Thus, the 11-year cycle is a process involving both solar hemispheres, although the development of the cycle is not entirely identical in both, and is not altogether synchronous.

Section 8. Certain Problems Involved in Solar Rotation and the Distribution of Solar Activity With Heliographic Longitude

Although the data in this section have been assigned by the author to this chapter, which discusses the cyclic character of solar activity, the phenomena which will be discussed here reveal no obvious dependence

on the 11-year or other solar cycles, and in a number of cases the problem involved is the finding of such a dependence.

The preceding section discussed problems associated with the distribution of solar activity by heliographic latitude. The principal purpose of this section is to present certain facts concerning the distribution of solar activity by heliographic longitude. However, since these problems are associated closely with the problem of solar rotation, it becomes necessary to discuss briefly the peculiarities of this phenomenon. In addition, in connection with the possibility of cyclic variations of the principal parameters characterizing solar rotation, it will be appropriate to briefly discuss variations in solar diameter, for which certain data indicate manifestations of the solar cycle. It is of course superfluous to note that the rate of rotation and radius of the sun can be related somehow to the law of conservation of angular momentum.

It is well known that the sun does not rotate as a solid body. The angular velocity of its rotation is a function of heliographic latitude. There are data indicating that this peculiarity of solar rotation was known by Galileo (Ref. 133), but it was established quite precisely by Carrington in the last century (Ref. 5). The studies of numerous observers along these lines have usually been made using empirical formulas describing this dependence of angular velocity on heliographic latitude. The most familiar and widely used formula at the present time is the Faye formula, having the form

$$\xi = a - b\sin^2\varphi, \tag{1.27}$$

where ξ is angular velocity, usually in degrees per day (however, for theoretical purposes it can be expressed conveniently in radians/second; the conversion is made using the relation: $1^\circ/\text{day} = 0.202\cdot 10^{-6}$ radian/second; φ is heliographic latitude; a and b are constants for the particular formula. In other formulas the sine is replaced by the cosine and the exponent 2 for sin φ is replaced by another power of the sine or cosine, but none of these formulas have appreciable advantages over the Faye formula; on the contrary, they give poorer results. It makes sense to replace the sine by the cosine only when latitude is replaced by polar distance, which sometimes is used in theoretical investigations (Ref. 134). Another problem is the numerical coefficients. In actuality these coefficients are different for the different indices from which the angular velocity of solar rotation is determined. Even with respect to the same index, such as for sunspots, the numerical values in formula (1.27) can differ if different series of observations are used; thus, the coefficient b from the Greenwich series of coordinates of sunspots for the period 1878-1923 is equal to 2.60, whereas from the series for 1924-1933 (same observatory), the value is 3.00.

This circumstance some time ago suggested the presence of a definite cycle in the values of the coefficients in formula (1.27). In actuality, the situation here is more complex; the large discrepancies in the values of the coefficients a and b for different epochs more likely are caused by systematic errors.

There are two principal methods for determining the rate of solar rotation; from the movements of visible details, caused by rotation, such as the movement of spots, faculas, etc., and spectroscopic observations. The latter make it possible to observe rotation at high latitudes, as much as 80° in the direction of the pole. In the high latitudes, rotation also can be determined from such formations as prominences, and studies recently appearing describe determination of the angular velocity of solar rotation at extremely high latitudes from observations of polar faculas (Refs. 135, 5). It should be noted that the spectroscopic method, despite all its attractiveness, can be burdened by serious systematic errors (Ref. 5).

Certain authors note that high-latitude formations (faculas and filaments) reveal rotation which cannot be described by formula (1.27), regardless of what values for a and b are used. In such cases it is necessary to use an expression in which the sine is replaced by the angle itself (Ref. 136). As an example we will cite a formula derived by Bruzek for the angular velocity of rotation of polar filaments

$$\xi = 13.36^{\circ} - 0.065^{\circ} \, (\varphi^{\circ} - 45^{\circ}),$$

correct in the latitude range $45^{\circ} \leqslant \varphi \leqslant 80^{\circ}$.

We note also that in certain cases it makes sense to introduce into the right-hand side of the Faye formula an additional term of the form $c \sin^4 \varphi$ (Ref. 5). However, such a supplementation of the Faye formula is most interesting in that it is in agreement with theoretical concepts concerning the law of solar rotation.

Recently considerable attention was given to the change in the rate of solar rotation with height, and the change with height of the dependence of the rate of rotation on heliographic latitude. In general, most investigators have concluded that the rate of rotation increases somewhat with height in the solar atmosphere and the dependence on latitude decreases. However, definite doubts have been expressed recently relative to these conclusions. Table 15, compiled using relatively old dats, is indicative.

The table shows that even if we ignore the fact that the accuracy of determination of the solar rotation rate from faculas and prominences is far lower than from spots, and that small deviations of the values a

and b for faculas and prominences from the corresponding spot values should not be ascribed too great importance, the dependence of the angular velocity of rotation on height is by no means of a simple character. As shown by the table, faculas have a larger value a (equatorial velocity) than spots, but prominences have almost the same value a as spots. The value b (equatorial acceleration) of faculas is higher, and of prominences appreciably less than for spots. However, on the basis of spectroscopic data, Aslapov (Ref. 137) quite recently was able to show that in the upper layers of the photosphere, angular velocity appears actually to increase somewhat with height, but as yet it is doubtful that this result actually reflects an increase of angular velocity. The investigations of Newton at Greenwich gave the results for the 11-year cycles Nos. 12-16 (Ref. 5) shown in Table 16.

The following expression for the Faye formula was obtained from the five cycles considered:

$$\xi = 14^{\circ}\cdot 38 - 2.77^{\circ} \sin^2\varphi \; .$$

Other, earlier determinations, based on fewer data, gave clues in certain cases to variations of the rotation rate as related to the solar cycle. An indicative result in this respect was that obtained by St. John, based on spectroscopic determinations of the rotation rate on the basis of data from Mount Wilson Observatory for 1900-1934. This result has been shown in Figure 19. By averaging the points, it is possible to obtain a certain clue to velocity minima in 1912 and about 1930, i.e., a slight hint of the possibility of the existence of a 22-year cycle. However, these data are too limited and not sufficiently convincing for drawing final conclusions on this problem. In addition, it must be mentioned that these are spectroscopic data with their inherent systematic errors. It is true that the principal error in such cases—scattered light in the instrument—has been taken into account here (Ref. 5), but we must agree with the authors, who warn against excessive trust in spectroscopic series for determination of the rate of solar rotation.

This same figure not only has a solid line denoting the linear velocity of solar rotation (in km/sec), but also a curve (2) representing change in the equatorial radius of the sun during these same years on the basis of meridian observations at the Monte Mario (Rome) Observatory. This value, given in seconds of arc, averages 961.30 sec and has an amplitude of about 0.9 sec. In Figure 19, seconds of arc have been replaced by an arbitrary scale linearly related to seconds of arc. The problem of variations of the solar diameter has a long history. The greatest attention to this problem has been given by Italian, and to a lesser degree, by German and British investigators. The clearest results have been obtained by the Italian astronomers Secchi, Rosa, Respighi, Armellini and Cimino.

Table 15

Object	Series of years	Place of observation	Coefficients of Faye formula	
			a	b
spots	1878-1923	Greenwich	14.43°	2.13°
faculas	1888-1923	Greenwich	14.54	2.81
prominences	1919-1930	Meudon	14.46	1.94

Table 16

No. of cycle	Years	Form of Faye formula
12	1878-1888	$\xi = 14.36^\circ - 2.5^\circ \sin^2\varphi$
13	1889-1899	$\xi = 14.39^\circ - 3.0^\circ \sin^2\varphi$
14	1900-1913	$\xi = 14.39^\circ - 2.8^\circ \sin^2\varphi$
15	1914-1923	$\xi = 14.39^\circ - 2.6^\circ \sin^2\varphi$
16	1924-1933	$\xi = 14.37^\circ - 3.0^\circ \sin^2\varphi$

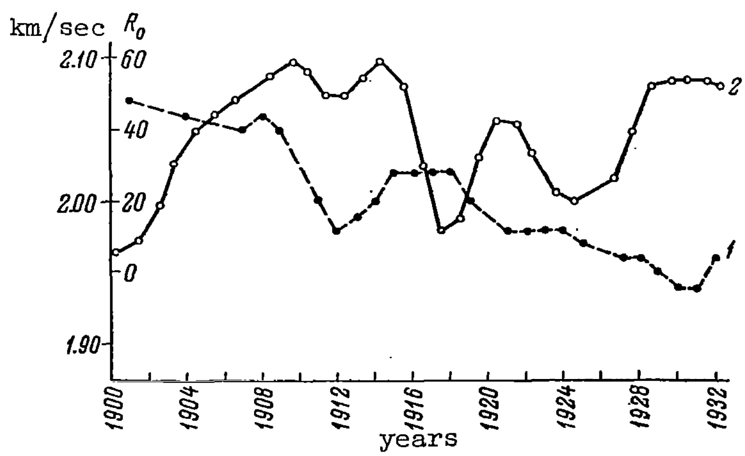


Figure 19. Comparison of variations of the rate of solar rotation and variations of solar radius; 1, solar rotation velocity; 2, variations in solar radius

We shall now discuss the results obtained by Cimino and his associates in Refs. 138 and 139. They were able to demonstrate that variations of the equatorial diameter can be represented as a principal variation, with a period of about 22 years, and a number of secondary variations, with periods of 6-1/2-8-1/2, 11 and 15 years. The maximum values of the diameter lag somewhat relative to the epochs of the minima of odd 11-year cycles: Nos. 13, 15 and apparently No. 17. These and somewhat earlier results were criticized, primarily by Greenwich Observatory specialists who considered the observed variations of the solar diameter to be the result of systematic errors. The annual variation of the solar diameter (after exclusion of variations associated with changes of the sun-earth distance) and certain other effects have been discussed in Ref. 140.

It is entirely natural that Italian investigators have attempted to confirm their results by citing series of observations of other observatories, especially Greenwich. Using these series it apparently has also been possible to determine the variations of the sun's polar diameter (Ref. 141), but the Greenwich astronomer Gething (Ref. 140) has denied the reality of such a result.

In this connection it is necessary to mention the conclusions of the German investigators Schur, Ambronn and Meyermann (Refs. 142, 143) from the processing of heliometer observations. Meyermann found that the amplitude of the variation of the diameter, 0.09 sec, was an entire order of magnitude less than that which was discovered by Italian investigators, that the cycle of variations is 11 years, not 22; and, the maximum diameter occurs at the maximum and the minimum at the minimum. However, the series processed by Meyermann was considerably shorter than the Italian series, and the small value of the amplitude in general forces us to doubt the reality of the variations determined with the Gottingen heliometer. However, there is no full assurance that the Italian series is not the product of systematic errors. Nevertheless, we are inclined to assume that they contain a grain of truth; we are led to this conclusion by Figure 19, which shows a comparison of the variations of the rate of rotation and the equatorial diameter. It can be seen from a comparison of the curves that there is a slight but definite negative correlation between the corresponding values. This is as it should be, if, as seems natural, the law of conservation of angular momentum is satisfied. Since Figure 19 shows the velocity curve, it can be concluded that the amplitude of angular velocity should be greater in that linear velocity decreases toward the epoch of the minimum of odd cycles despite an increase of diameter with an approach to this epoch. An unusual situation arises. Each of the two series separately is of low reliability, but their negative correlation suggests the reality of an opposite course of variation of these values. We cannot compute the coefficient of the negative correlation here because each of the series

is smoothed, and diameter variations are twice smoothed. Thereby, the negative correlation suggests the reality of each of the series separately, since the opposite variation of these values is, as might be expected, on the basis of the direct concepts of mechanics.

We shall now consider the nonuniform distribution of solar activity by heliographic longitude. This phenomenon was discovered for the first time by Wolfer in 1889 while investigating spots, faculas and prominences (Ref. 144). Wolfer also noted that there apparently are two longitudinal concentration zones of activity on the sun which are situated approximately 180° apart. The longitudes predominantly occupied by active formations, or active longitudes, as they are now usually called, are discovered most easily from faculas. In certain years of the cycle, the duality of active longitudes is expressed more clearly than in others. Wolfer noted a shift of active longitudes with time and related this to a change of the latitude of the zone of activity in the cycle, that is, with a change of the rotation period in accordance with the law of solar rotation. Of the numerous later studies devoted to this problem, we will discuss primarily the investigations of Maunder, Deslandres, Losh and U. Becker. As will be described later in considerable detail, active longitudes are observed not only in solar phenomena, but to a still greater degree in solar-induced geophysical phenomena. The study made by Maunder (Ref. 145) in part discusses the active longitudes of geophysical phenomena, specifically magnetic storms, and we will not touch upon this part of his work.

The part of Maunder's paper which applies to purely solar phenomena is of methodological interest because in it he initiated the use of 27-day diagrams in the study of solar-activity statistics. These diagrams later came into very wide use, especially for geophysical purposes, in the investigation of solar-induced phenomena. In his investigation of spots for the period 1891-1902, Maunder noted that it is possible to select a sidereal period of solar rotation in which the active longitudes are the most clearly expressed and the most stable. He worked initially with a mean sidereal period of 25.38 days, and then changed to a period of 25.09 days, which was more suitable. However, Maunder was not able to establish the true period of rotation of active longitudes. He stated only that the taking into account of Faye's law, that is, the use at each heliographic latitude of that period of rotation which corresponds to it according to the formula $\xi = a - b \sin^2 \varphi$, does not make it possible to establish the real period of rotation of active longitudes. Maunder's work was followed by a whole series of investigations devoted to active longitudes of both purely solar and purely geophysical phenomena associated with solar activity. We will not discuss each of them because a detailed review of the problem is contained in a book by M. S. Eygenson (Ref. 75).

We will discuss the work of Losh, published in 1938. This author studied Wolf numbers for 1903-1937 and spot areas for 1916-1934. A summing for 10 solar rotations in 20° bands of longitude revealed that activity is distributed nonuniformly by longitude. There are passive as well as active longitudes. The extrema of the curves have a tendency to lie 180° apart. It was possible to trace the active longitudes for 200 solar rotations (more than one 11-year cycle) and even more, but an active longitude sometimes becomes definitely passive and vice versa. In all the investigations of the 11-year cycles it was characteristic that the active and passive longitudes were separated by 180°. In general, there is a constancy of active longitudes from cycle to cycle, but according to the Wolf number index there is a tendency to an alternation from cycle to cycle of the active and passive longitudes in a particular longitude interval and it can be said with some degree of assurance that there is a 22-year activity cycle in a particular longitude interval. Latest data do not confirm this point of view and also cast doubt on the tendency to an antipodal character of purely solar phenomena. Antipodal characteristics for hemispheres can be considered with a somewhat greater degree of assurance. For example, if a particular longitude interval was active in a particular 11-year cycle in the solar northern hemisphere, there will be an active interval in the southern hemisphere in this same epoch at a distance of 180° from the other. It has not been possible to discover a regular pattern of movement of active longitudes.

With respect to antipodal behavior, it is impossible to pass over an investigation made by Waldmeier (Ref. 5) in which he established the tendency of spot groups to concentrate simultaneously in the following latitude-longitude zones:

$$\begin{aligned} \varphi &= b, & \lambda &= 1 \\ \varphi &= -b, & \lambda &= 1 \\ \varphi &= b, & \lambda &= 1 + 180^{\circ} \\ \varphi &= -b, & \lambda &= 1 + 180^{\circ}. \end{aligned}$$

Before discussing other, later investigations of active longitudes, it is necessary to discuss briefly the methods used to detect them. The most frequent and widespread method used is the so-called "27-day" solar calendars introduced by Maunder, which are very commonly employed in the investigation of the frequency of recurrence of geophysical phenomena. In such work certain geophysicists employ lines of 27 days exactly, ignoring the known deviation of the actual period of synodic solar rotation from this figure (Ref. 146). Needless to say, it is more correct to take into account the deviations of the true period of solar rotation from 27 days. The calendars usually are constructed in the Carrington longitude system. In this case the synodic period cannot be assumed to be 27.33, 27.25 or, in general, any other value constant for the entire year. If, as assumed in the Carrington system, it

is possible to speak of a rigorous sidereal period, since it applies here to a definite heliographic latitude (15°), the synodic correction nevertheless cannot be constant for the entire year because it is dependent on the sun-earth distance. In constructing solar calendars it, therefore, is necessary to use the commencement dates of the Carrington rotations cited in astronomical yearbooks. If the commencement time of a particular rotation, given in the yearbooks with an accuracy to 0.01 day, falls in the afternoon of a particular day, that is, the fractional part of the corresponding day is less than 0.5, it is assumed that the particular solar rotation begins on the following day.

After such a calendar of solar phenomena has been constructed for a sufficiently large series of solar rotations, the values of the indices plotted in it are summed by columns, each of which embraces a longitude interval of about 13°. As a result it is possible to construct a curve with Carrington longitudes plotted along one axis and the total value of the particular solar index in the corresponding longitude interval plotted along the other. We will refrain from calling such a curve a "curve of distribution of solar activity by longitudes", since the distribution curve has a somewhat different meaning. It gives the frequency of recurrence of the value of the argument itself, and in this particular case the argument is longitude, at the same time that the dimensionality of the ordinates corresponds to the dimensionality of that index used in a construction of the calendar.

In old investigations it was customary upon obtaining summary curves to determine the point and proceed to a discussion of the results. However, it is absolutely necessary to clarify the nonrandom character of these results before drawing any conclusions from them. In recent decades many statistical criteria have been proposed for investigation of the nonrandom character of the distributions obtained from solar calendars. For example, it has been pointed out that it is possible to use the B. P. Veynberg criterion, and later the Gleissberg criterion (Ref. 147), but there is no doubt but that it is more appropriate to use for this purpose one of the statistical tests of fit (X^2 or λ), by means of which the obtained distribution is compared with a "zero hypothesis", characterizing a uniform distribution of a particular index relative to solar longitude. We used such a method for clarification of the nonrandom character of the summary curve of the solar calendar of spot groups with an area of >100 millionths of a hemisphere for 1933-1940, and a curve of this same kind for flares during 1935-1940 (Ref. 148). For the purpose of consolidating the class of the argument, in many cases the data for "three-day" periods in the solar calendar are summed yielding 9 instead of 27 or 28 values of the argument (the ninth is obtained from the columns for the 28th day and obviously contains lesser values). But this method cannot be used in all cases. For example, if the problem is to find the spectrum of active longitudes, it

is clear that this consolidation into three-day periods can distort the picture, and in this case it is necessary to proceed on the basis of all possible values of the argument.

In 1933, by using data on flocculi for 1919-1927, Deslandres (Ref. 149) revealed that these phenomena are concentrated in longitude intervals which are multiples of 15°. It is impossible to draw such a conclusion when the results are combined into three-day periods. It is true that the results of Deslandres have been subjected to criticism by Salet (Ref. 150) and by certain other authors, so that at present there is no assurance of such a separation of active longitudes.

The data which we investigated were broken down into three-day periods; nine classes of the argument were considered. Preference must be given to the A. N. Kolmogorov λ criterion, since when using the Pearson X^2 test, the investigator meets with a certain ambiguity relative to the number of degrees of freedom. The results of the study revealed that, in the case of spots, the probability of the derived distribution relative to heliographic longitude being random is 0.04 (that is, falls within the limits of the confidence coefficient 0.05, whereas for flares it is still less--several ten-thousandths. Almost simultaneously with our application of the tests of good fit to the 27-day calendar, Brooks used it for several other problems, which although different were similar in a mathematical sense; these problems were also from the field of the effect of solar activity on geophysical phenomena (Ref. 151). The procedures used by Brooks were more nearly perfect than those of the author; he not only found, by use of the X^2 test, that the curves which he investigated were real, but also used the t test to check the reality of the determined extrema of these curves. This combined mathematical-statistical check was named the "n-method" by Brooks. We feel that the study of active solar longitudes has entered a new stage in which tests of good fit and other methods of mathematical statistics can be used for checking the reality of both curves as a whole and their individual extrema. Because of the lack of these tests in older studies, the latter cannot be considered convincing. Without objective indices of the reality of the determined active longitudes, it is possible to identify erroneously certain longitudes as such, and permit actual active longitudes to escape notice.

An interesting illustrative example is one of the studies of U. Becker (Ref. 152); in this investigation he discusses, in particular, the longitude distribution of active centers. Although this is not exactly what is meant by the term "active longitudes", the statistical methods for evaluating the nonrandom character of the derived distributions remain the same. U. Becker cites a map of the distribution of active centers, or foci, and asserts without proof that their distribution relative to heliographic longitude is random. The author's

analysis revealed that in actuality the probability of a random character of the longitude distribution of these centers is very small.

Yu. I. Vitinskiy recently proposed a completely different method for the detection of active longitudes (Ref. 153). This method not only makes it possible to detect active longitudes, but also to clarify the extent of extremely long-lived centers along the meridian. We have here a new principle for construction of solar synoptic charts. After breaking the entire photosphere down into a series of rectangles by heliographic coordinates, Vitinskiy draws isolines of the probability of the appearance of a particular spot area in each of these grid squares. Each such solar map is something like a composite map for a long interval of time; Vitinskiy has used the 11-year cycle as such an interval (Figure 20). The active centers, especially those closely tied to particular longitudes, stand out very clearly; they move from one 11-year cycle to the next, but the movement is insignificant. This work of Yu. I. Vitinskiy confirms graphically what already follows from the work of Losh (Ref. 154); that is, that the internal factor responsible for sunspots is not subject to the influence of differential solar rotation.

An attractive feature of the Vitinskiy "isoline method" is its graphic character and the possibility of immediately obtaining the two-dimensional distribution of active centers of very long lifetime in relation to heliographic coordinates. It is clear, however, that the mathematical-statistical soundness of the results of such research methods still has not been demonstrated, although in principle this can be done.

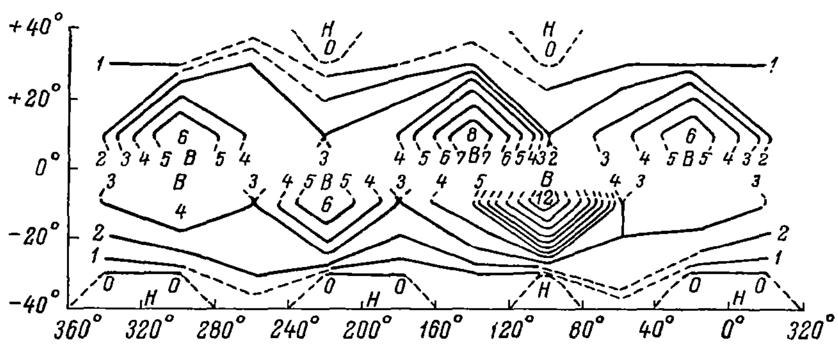


Figure 20. Distribution of long-lived active centers on the surface of the sun (according to Vitinskiy)

The details of investigation of active solar longitudes in geophysical phenomena will not be discussed here, but will be dealt with in part in Chapters 4 and 5. However, it must be that active longitudes in a number of cases are detected more clearly and are more stable relative to geophysical indices than relative to solar activity indices. Maunder used a solar calendar for determining the longitude (in the sense of heliographic longitude) distribution of magnetic storms. One, and in certain cases two, active longitudes were found. This result later was confirmed by Cortie (Ref. 155).

Bartels' investigation (Ref. 146), using data on magnetic storms for the period 1906-1931 led to the discovery of M-regions concentrated at the active longitudes. The antipodal character of active longitudes, discovered using geomagnetic data, is expressed more sharply than for purely solar indices, but it is manifested still more sharply at well-defined active longitudes which can be discovered on the basis of meteorological phenomena. In this connection we should mention the numerous studies made at Tashkent by M. S. Zhukov, P. P. Predtechenskiy and K. P. Brodovitskiy (Ref. 156). Two antipodal active longitudes, for example, are reflected in the discharge of the Syr-Darya River during the period of 1898-1912, and atmospheric pressure at Tashkent in 1936. We also can mention the longitude distribution of the dates of Arctic intrusions into a sector of the European natural synoptic region during the period of 1933-1937 (Ref. 157); this will be discussed in greater detail in Chapter 5.

These considerations make it possible to cite the following results of current concepts concerning active regions.

1. This phenomenon is entirely real, not apparent.

2. The stability of the longitudes is considerable; in the course of several cycles, certain active centers move not more than 40-60° in longitude. Taking into account that the extent of active longitudes falls in the range from 40 to 150° (to be sure, the lower limit is related to the method for detection of active longitude), such movements should be considered small.

3. The stability of active longitudes indicates a rigid rotation of the sun at those depths where the factor responsible for solar activity is concentrated.

4. There is a certain tendency to an antipodal character of active longitudes; a particular longitude rather frequently corresponds to another which is 180° away. However, in such cases the second longitude can differ from an ordinary active longitude not only in its activity, but also by the fact it is "passive", i.e., the activity is systematically not above the norm. A combination of these effects, with

the concentration of active longitudes from time to time in one solar hemisphere (northern or southern), rather than in both, leads to a complex picture whose most probable pattern has been determined by Waldmeier (see above).

5. A study of active longitudes by the use of the sunspot index reveals, as shown in Ref. 158, that the principal population of these longitudes consists of repeating groups with a considerable area (more than 500 millionths of a hemisphere; such groups are twice as numerous at active longitudes than at passive longitudes).

6. In a number of geophysical phenomena it also is possible to detect active longitudes, and to detect them more clearly than in many solar indices.

7. The antipodal character of active longitudes, as indicated by geophysical indices, is expressed more clearly, observed more frequently, and is more stable than is indicated by solar indices.

8. In an investigation of the longitude distribution of centers of geomagnetic activity, it is possible to note that in a number of cases they are located at distances which are multiples of 15° from one another, and sometimes are 40° apart. The simplest variant of the first case suggests two antipodal active longitudes, i.e., separation by an angular distance of 180°. The simplest variant of the second case suggests three active longitudes spaced at angular distances of 120° from one another (Refs. 150, 156).

9. There is no obvious relationship between active longitudes determined from either geophysical or solar indices, and the solar cycle. Such longitudes can persist for several 11-year cycles. The tendency to a 22-year cycle noted by certain authors for the appearance of active longitudes at similar meridians cannot be considered proven.

In conclusion of this section and the chapter as a whole, natural movements of solar formations relative to heliographic longitude will be discussed. Here, as already mentioned at the beginning of this section, there is no relationship to the solar cycle, and in this respect movements in longitude (to be more precise, along the solar parallel), also differ from movement along the solar meridian associated with the 11-year cycle and possibly with the 22-year cycle, as active longitudes of a zone of activity. The relationship of the latter with the 11-year cycle is the most direct, but as we have just seen, no such relationship has been established in the case of active longitudes. Natural movements along a parallel are detected by the observer as deviations from the law of rotation characteristic of a particular latitude. Investigations of this problem unanimously have associated natural movements in longitude with diverging movements in sunspot groups. Such a

phenomenon is noted most clearly in the first days of existence of a group, that is, when it still is forming. By the beginning of the breakup of a group, such diverging movements virtually cease, and with their cessation there is an end to all movements in longitude. At the beginning of the existence of a group, the rate of movement in longitude is about 0.6°/day, which corresponds to 7,000 km/day. This is approximately an order of magnitude greater than the rate of movement in latitude (Ref. 5). Bipolar group movement is usually characterized by the leading spot moving westward with a speed close to the abovementioned 0.6°, although a speed of 1°/day is sometimes attained.

Having moved approximately 10° westward by the end of the period of its development, the leading spot then begins to move slowly in the opposite direction during the time of decay of the group; that is, it moves eastward at a speed of 0.05-0.10°/day. By the time of the final destruction of the group, the leading spot is located at approximately the same place as it was initially. At the beginning of its development the following spot moves eastward at a speed of about 0.2°/day. However, the movement of the following spot ceases by the time of maximum development of the group. However, in certain cases deviations from this pattern are observed. For example, Waldmeier investigated a group having a considerable (30°) heliographic latitude in which the leading spot first moved 7° westward during the development of the group, and later, as a result of eastward movement, not only reached its former position, but moved 6° beyond. It has been found that the most rapid movement of a spot along the meridian is observed when the speed along the parallel becomes equal to zero.

The natural movements of other formations in longitude have been poorly studied and are apparently insignificant. For example, in the process of the development of active centers, filaments change their orientation, first extending along the parallel and then moving toward the pole. However, the change in orientation of a filament can be interpreted as curvature of a meridian, associated with differential solar rotation (Ref. 159).

Such are certain of the empirical data on the solar cycle and related problems. Certain problems have been purposely passed over in this chapter, leaving them until corresponding results are not required due to the development of definite theoretical concepts.

CHAPTER 2
STRUCTURE AND DYNAMICS OF THE SUBPHOTOSPHERIC LAYERS OF THE SUN

Section 1. General Information on Physical Conditions in the Subphotospheric Layers of the Sun

Gas Kinetic Characteristics

In the preceding chapter we discussed, with no pretense to exhaustiveness, but on a sufficiently broad basis, data applying to the field of solar activity, and to its cycles in particular. We shall now proceed to a theoretical interpretation of those facts (or at least some of them) which were discussed in Chapter 1. Before presenting the various hypotheses advanced for the explanation of these phenomena, including a number of concepts expounded in the author's own studies and those presented for the first time in this book, it is necessary to present certain general information on those layers of the sun in which activity phenomena are generated and which in all probability are responsible for the cyclic character of solar activity. Considerable development of activity occurs in the outer layers in the solar atmosphere; however, it is the inner layers (Ref. 1) which are responsible for the cyclic character.

The subphotospheric layers of the sun may be represented formally as extending from the lower boundary of the photosphere to the innermost regions. The methods for investigating these layers, therefore, are the methods used in the theory of the internal structure of stars. However, in stars it is the layers close to the photosphere that present specific difficulties. The approximation of a polytropic sphere and computations of characteristics using Emden function are unsuitable there, and it is necessary to formulate the problem of matching the solution of the problem of computation of a model for these layers with that obtained from an approximation of a polytropic sphere for deeper layers.

Recently, specialists on the internal structure of stars have stated directly that in the study of the peripheral layers, it is necessary to consider the characteristics of turbulence, thereby introducing a strong element of inaccuracy into the theory (Ref. 2). It is possible that it would be correct to distinguish the physics of the peripheral part of the sun's inner layers from the general theory of

the internal structure of stars, in the same way that the physics of the earth's ionosphere is a special division of aeronomy.

One of the specific difficulties in developing a model of the sun for the peripheral part of its inner layers is that in the computations it is impossible to neglect changes with depth of temperature, pressure, density and other characteristics of such properties of the subphotospheric solar gas as viscosity, heat conductivity, electrical conductivity, etc. These properties themselves are dependent on temperature and the other characteristics forming the content of the model. This circumstance inevitably leads to the necessity of using the method of successive approximations, but care must be exercised. In most cases it would be desirable to use as a first approximation a model closest to that which would be obtained after taking into account those factors which previously have never been considered or have been considered inadequately. However, it can happen that with an inadequate, or especially with an incorrect, allowance for certain factors the initial approximation may be found to be worse than if these factors were not taken into account at all. We therefore feel it is rational to use as a first approximation a model which does not take into account those factors that introduce complications into the models of higher approximations. It may be more correct to call this a zero approximation.

The only correct method for the computation of the principal characteristics of stratification is that of successive approximations, or to be more exact, computation by intervals (Ref. 3). In order not to burden the reader with a somewhat abstract method of reasoning, we will use an example to illustrate what is involved. One of the not unimportant characteristics of the subphotospheric medium is the gas kinetic viscosity of the gas of which it consists. Such viscosity must be included in the computation in the determination of the velocity of convective flow or in an evaluation of the character of the convection, since viscosity is involved in the corresponding criteria. However, viscosity of a partially ionized gas is dependent on the degree of ionization. Therefore, it would seem that if we wish to determine viscosity at a particular level under the photosphere, we first would have to establish whether the solar gas there is completely ionized, and if it is not, to what degree it is ionized.

The physical characteristic, i.e., the degree of ionization, has, however, the most primary significance for a factor far more important than gas kinetic viscosity, allowance for which can be neglected, however, in this first approximation; the degree of ionization is decisive for stratification in general, because it is related directly to the adiabatic gradient in the gas of which the sun consists. For a clarification of stratification it is necessary to have the most precise data possible concerning ionization. For the determination of viscosity, however, considering that the degree of ionization is important, but

not of primary significance, it is possible to limit ourselves to a determination of ionization in a certain approximation. Since ionization in the subphotospheric layers is determined primarily by temperature, what we have said above means that in order to determine viscosity, it is possible to use approximate data on temperature distribution with depth. This is also true for other characteristics having a nature similar to viscosity, such as heat conductivity, electrical conductivity, etc.

Among the properties and characteristics of the subphotospheric layers to be discussed here are chemical composition and the degree of ionization at different depths (as is clear from the preceding discussion, initially it is sufficient to make an approximate estimate). This will be followed by a discussion of the previously mentioned gas kinetic characteristics: viscosity, heat conductivity and electrical conductivity. We will avoid the term "molecular" characteristics because we are dealing with an atomic, not a molecular gas.

Thereafter, we consider it necessary to discuss at some length the qualitative or semiqualitative aspects of the macrocharacteristics, that is, to proceed from gas kinetic to the hydrodynamic and magnetohydrodynamic properties of the subphotospheric gas. Only after such a general physical consideration of the medium with which we will be concerned for the most part in this and succeeding chapters will it be possible in the next section to continue to a brief historical review of concepts of the stratification of the subphotospheric layers and the present-day theories concerning this problem. We will in this manner prepare for a transition to the next section which deals with the problems involved in the dynamics of the layers with which we are concerned. However, even at this point it can be noted that with respect to dynamic factors the situation to a certain degree is similar to that of gas kinetic factors: the dynamic, especially hydrodynamic, factors are related closely to stratification, but in turn they themselves influence stratification. This circumstance also introduces certain difficulties into the study of this particular set of problems, but it is, nevertheless, thermodynamic conditions, i.e., stratification, which are of primary significance.

The first of the formulated problems—the chemical composition of the subphotospheric layers--can be solved quite simply. If the composition of the solar interior does not differ too greatly from the composition of the solar atmosphere, the difference between the chemical composition of the atmosphere and the directly adjacent subphotospheric layers should be still less.

With respect to the distribution of ionization with depth, in computing the gas kinetic parameters it is possible to use approximate values. As a zero approximation we will use the moded proposed by A. G.

Masevich and T. G. Volkonskaya (Ref. 4), since in this model the authors do not determine the outer convective zone. Therefore, when discussing gas kinetic characteristics we will not deal with the a priori existence of stratification causing the presence of an outer zone of convection.

Like all models used as a zero approximation, the Masevich-Volkonskaya model does not include layers situated directly under the photosphere. In order to include these layers we used the elementary distribution of pressure, temperature and density in the lower photosphere (Ref. 5) derived from solution of the photospheric problem, and made an interpolation of values in the transitional region between depths of 130 and 2,200 km. Depths were determined from $\tau = 1.00$. The following chemical composition was assumed in the Masevich-Volkonskaya model: X (hydrogen) = 0.995, Y (helium) = 0.003, Z (metals) = 0.002. Therefore, the atomic weight at depths greater than 2,200 km was assumed to equal 0.5, and the gas pressure was computed using this value (in this model, temperature and density are stated). In view of the interpolation, the gas-pressure values in the transitional region were found to be generally too low at each particular depth. As a result, the values of the electron pressure in the region close to the zone of total ionization were higher than is physically admissible. In such cases we used the extreme value of electron pressure for a hydrogen medium.

The determination of the gas kinetic characteristics, that is, viscosity, heat conductivity and electrical conductivity of the gas mixture, which constitutes the subphotospheric gas characterized by a changing degree of ionization with depth, is an exceedingly difficult problem. It therefore is entirely understandable that the scientists who have investigated this problem have attempted to simplify it. In a study by Oster (Ref. 6), representing one of the most thorough attempts to establish the values of the gas kinetic characteristics of the solar gas, formulas have been derived which make it possible to compute viscosity, heat conductivity and electrical conductivity for a completely neutral or a completely ionized gas.

For an ionized gas, Oster cites curves for all three gas kinetic characteristics with the argument of electron pressure and the parameter $\theta = 5{,}040/T$. This parameter is given in the range from 0.4 to 0.001, which corresponds to temperatures from 12,600 to $5 \cdot 10^{6}$ $^{\circ}$K. At the same time, the argument--electron pressure logarithm--has been given in the range from -2 to +5; the latter corresponds in the model that we have used to a depth of only several hundred kilometers under the photosphere, $\theta = 0.4$. The Oster curves, therefore, are applicable for the most part to the solar atmosphere, but here the conductivity computations are more complex than those which can be made solely on the basis of Oster's theoretical premises (Ref. 7).

For the purpose of determining the gas kinetic characteristics in the entire subphotospheric layer considered in our model (Table 17), we made a number of additional computations, using the formulas cited by Oster. The Oster formulas are applicable to the entire thickness insofar as viscosity and heat conductivity are concerned. The same cannot be said of electrical conductivity; the Oster formulas are suitable for the transitional zone where hydrogen is ionized partially, but the well-known Cowling formula (Ref. 8) is more suitable for the region of its complete ionization. The principal gas kinetic characteristics have been given in Table 17.

Characteristics Associated With the Radiation Field

In this section we will discuss the problem of radiative viscosity and the radiative heat conductivity of the subphotospheric layers. The following formula from Ref. 8 can be used for computation of the first of these characteristics:

$$\mu_r = \frac{4p'}{5\bar{\varkappa}\rho c}, \tag{2.1}$$

where p' is radiation pressure: $p^1 = 1/3aT^4 (a = 7.66\cdot10^{-15})$; c is the speed of light; $\bar{x}$ is the coefficient of continuous absorption (Rosseland coefficient); and ρ is density. The value $\bar{x}$ was taken from the same Masevich-Volkovskaya model; in the layers close to the photosphere we used data cited by Allen; for the intermediate regions we used graphic interpolation. Since $\bar{x}$ does not change too significantly in the subphotospheric layers, attaining large values only relatively close to the photosphere (zone of intense ionization of hydrogen), such an approximation is entirely adequate for the computation of radiative viscosity. The remaining characteristics were taken from Table 17.

Radiative heat conductivity was computed using a formula from Ref. 9:

$$K_r = \frac{\sigma T^3}{\varkappa\rho}, \tag{2.2}$$

where σ is the Stefan constant ($\sigma = 5.75\cdot10^{-5}$) and the remaining values are the same as in formula (2.1). The results of the computations are given in Table 18.

By comparing these results with the data in Table 17 we see that radiative viscosity at almost every specific depth somewhat exceeds the gas kinetic viscosity; at a depth of about 100,000 km under the photosphere this excess attains almost a complete order of magnitude. That the radiative heat conductivity exceeds gas kinetic heat conductivity by many orders of magnitude is well known, and explains the predominance of radiative transport as a form of energy transport.

It is interesting to compare the gas kinetic viscosity of solar gas with the molecular viscosity of certain terrestrial substances. For example, at normal pressure and temperature air has a viscosity of $1.72 \cdot 10^{-4}$ g/cm·sec, which corresponds to the gas kinetic viscosity of the solar gas at a depth of 4,000 to 5,000 km below the photosphere, or radiative viscosity at a depth of about 3,000 km. At a depth of 70,000 km the viscosity of the solar gas is equal to the viscosity of strong sulfuric acid. The radiative viscosity on the sun at a depth of about 100,000 km is approximately the same as the molecular viscosity of castor oil. The molecular heat conductivity of the air, which under normal conditions is $2.34 \cdot 10^2$ ergs/cm·sec·°K, does not correspond to the gas kinetic heat conductivity of the solar gas in the subphotospheric layers and is two orders of magnitude lower than its heat conductivity, even at a depth of 500 km below the photosphere. This also is completely understandable if we take into account the high temperature of the gas on the sun and the presence there of free electrons as the carriers of thermal energy.

Table 17

Depth underneath the photosphere z, cm	Absolute temperature T°, K	Density ρ $g \cdot cm^{-3}$	Gas pressure, p_g, $dynes \cdot cm^{-2}$	Dynamic viscosity μ, $g \cdot cm^{-1} \cdot sec^{-1}$	Heat conductivity k, $g \cdot cm^{-1} \cdot sec^{-1}$/deg	Electrical conductivity
$5 \cdot 10^7$	$1.26 \cdot 10^4$	$1.12 \cdot 10^{-6}$	$2.34 \cdot 10^6$	$9.65 \cdot 10^{-6}$	$2.08 \cdot 10^5$	$3.02 \cdot 10^{-8}$
$1 \cdot 10^8$	1.58	1.58	4.15	$1.87 \cdot 10^{-5}$	4.15	7.55
2	1.66	2.51	6.31	2.22	4.97	7.83
3	1.74	3.02	8.73	2.53	5.78	8.12
4	2.51	3.72	$1.58 \cdot 10^7$	6.85	$1.62 \cdot 10^6$	(9.62)
5	3.16	4.47	2.24	$1.24 \cdot 10^{-4}$	3.02	$1.12 \cdot 10^{-7}$
6	3.55	5.62	3.32	1.68	4.16	1.33
7	3.98	6.31	3.98	2.23	5.57	1.59
8	4.47	6.61	4.47	3.04	7.40	1.86
9	5.01	7.08	5.90	3.95	9.79	2.24
$1 \cdot 10^9$	5.25	7.59	6.62	4.46	$1.10 \cdot 10^7$	2.40
2	8.51	$2.00 \cdot 10^5$	$2.83 \cdot 10^8$	$1.52 \cdot 10^{-3}$	3.65	4.96
3	$1.26 \cdot 10^5$	5.62	$1.18 \cdot 10^9$	3.93	9.64	8.94
4	1.51	$1.12 \cdot 10^{-4}$	2.51	6.30	$1.52 \cdot 10^8$	$1.17 \cdot 10^{-6}$
5	2.51	2.00	7.94	$2.10 \cdot 10^{-2}$	5.43	2.52
6	2.82	3.98	$2.00 \cdot 10^{10}$	2.82	6.56	3.00
7	6.31	10^{-3}	$1.05 \cdot 10^{11}$	$1.85 \cdot 10^{-1}$	$4.04 \cdot 10^9$	$1 \cdot 10^{-5}$
8	8.91	$1.26 \cdot 10^{-3}$	1.87	4.10	8.77	$1.68 \cdot 10^{-5}$
9	9.97	1.52	2.52	5.30	$1.14 \cdot 10^{10}$	1.99
$1 \cdot 10^{10}$	$1.12 \cdot 10^6$	1.78	3.31	7.00	1.48	3.38

Table 18

z	μ_r	K_r
$5 \cdot 10^7$	$5.47 \cdot 10^{-6}$	$3.64 \cdot 10^{11}$
$1 \cdot 10^8$	$4.48 \cdot 10^{-5}$	$2.38 \cdot 10^{12}$
2	$1.06 \cdot 10^{-4}$	5.38
3	1.47	7.12
4	5.55	7.38
5	$1.21 \cdot 10^{-3}$	$3.24 \cdot 10^{13}$
6	1.57	3.72
7	2.26	4.78
8	3.44	6.47
9	5.15	8.67
$1 \cdot 10^9$	5.95	9.55
2	$1.67 \cdot 10^{-2}$	$1.66 \cdot 10^{14}$
$3 \cdot 10^9$	3.07	2.05
4	3.47	(2.09)
5	$1.60 \cdot 10^{-1}$	2.13
6	$(1.20 \cdot 10^0)$	4.01
7	1.37	$1.85 \cdot 10^{15}$
8	4.75	4.48
9	6.49	5.49
$1 \cdot 10^{10}$	9.10	$6.88 \cdot 10^{15}$

Macrocharacteristics of Subphotospheric Layers

Among the macrocharacteristics of the subphotospheric layers are turbulence, convection, various kinds of wave motions, and the properties of magnetic fields situated at depth.

In this section we will discuss the problems of turbulence and magnetic fields. The convection problem will be discussed in the next section.

We will recall the principal properties of turbulent motions with respect to astrophysical conditions. Under these conditions it is quite typical for there to be development of chaotic vortical motions of gas at a large scale. These fluxes have a tendency to develop into vortices of smaller scale (eddies). This physical property also is reflected in the form of the corresponding hydrodynamic equation; the term containing the Euler operator in the expression for the individual derivative in the equation of motion makes this equation nonlinear.

If velocity consists, therefore, of velocities of motion of at least two scales, the nonlinearity leads to a relationship between motions of different scales (Ref. 10). Thus, the energy of motion of one scale passes into the energy of motion of the other, and considering the second law of thermodynamics, it is more probable that there will be

a transition from motion of a higher order to a lower order. The limit to which the motion will go from the larger to the smaller scales is dependent on gas kinetic viscosity; if it is great, the transition of energy of motion into thermal energy occurs when the principal motion has not degenerated into very small pulsations. In any case, more or less small pulsations die out due to viscosity, and the energy of motion is converted into thermal energy.

Thus, an important property of turbulent motion is a flux of energy from the principal motion to small pulsations; the maintenance of turbulent motion is possible only when the principal motion receives the same quantity of energy as is carried by the abovementioned flux from the higher to the lower scales. In turbulent motion there always is an array of turbulent elements or elementary eddies. The following is what is meant by a turbulent element: if two points are arbitrarily taken in a liquid (gas), separated from one another by a small distance, the motions of these points are interrelated closely. By each time selecting two points which are increasingly distant from one another, eventually it is possible to obtain a pair in which the motions of the components will in no way be related to one another. The distance between the components of such a pair is called the dimension of the turbulent element. The distance which the turbulent element travels before losing its "individuality", that is, before merging with the surrounding medium, is called the mixing length.

Since the study of the properties of turbulent motions is conceivable only as a statistical investigation, that is, by replacement of individual objects by their mean characteristics, we discuss, in particular, the mean mixing length. Each eddy or turbulent element can be assigned a definite linear dimension l, and characterized by the wave number k, equal to $2\pi/L$. One of the important values in turbulence theory is $\rho F(k)dk$, representing the energy of vortices in a unit volume and those whose wave numbers range from k to k + dk. The function F(k) determines the turbulence spectrum. With a decay of the principal motion (principal flux) the motion nearest in the turbulence hierarchy to the principal motion retains features of the principal motion. The developing pulsations are different in dependence on their coordinates. For the next member of the hierarchy in scale, this dependence already is expressed more weakly, and for the next, still more weakly, etc. Motion in the smaller pulsations becomes isotropic, and it is here that the most satisfactory theory of turbulence has been developed of the many turbulence theories.

However, this does not exclude the nonisotropic character of "parental" pulsations in relation to the initial motion and the developed theoretical model does not become unsatisfactory since all of the pulsations are carried off by the initial flux. This situation is of appreciable importance because certain authors erroneously assume

that the presence of an ordered principal flux prevents the possibility of use of certain of the basic principles of turbulence theory, especially for determination of the coefficient of turbulent viscosity (Ref. 11). The form of the function F(k) has been investigated many times. The best expression for this function is the so-called Kolmogorov-Obukhov "2/3 law", according to which energy is proportional to wave number with an exponent of 2/3 (Refs. 12, 13). However, it is possible to cite examples in which this law is not satisfied. Such examples can be found in astrophysics (photospheric turbulence, Ref. 14), and in meteorology (large-scale motion) for which a "law of the first power" (Ref. 15) is closer to the truth.

An important result of the theory of stationary isotropic turbulence is that beginning at a certain scale the characteristics of pulsations are dependent only on the principal, decisive parameters of motion. In the case of the smallest spatial scales such parameters are the flux of energy coming from motion of a higher scale and kinematic viscosity $\nu = \mu / \rho$.

It is the energy flux which is of the greatest importance for motion at more considerable scales. It should be noted that these considerations are correct for an incompressible liquid. The most widely used criterion making it possible to judge when laminar flow undergoes transition into turbulent flow is the Reynolds criterion

$$\mathrm{Re} = \frac{vL}{\nu}, \tag{2.3}$$

and it can be assumed with sufficient accuracy that turbulent motion develops when Re $> 1{,}000$. The meaning of this criterion is obvious if we take the preceding comments on the role of viscosity as a factor either suppressing motion at a particular scale and changing its energy into thermal energy, or at lesser values, permitting transition to motion at a lower scale.

A comparison of the inertia term in the equation of motion, whose order of magnitude is v/l, and the term which takes viscosity into account, whose order of magnitude is v/l^2, directly gives the Reynolds criterion. Recalling what was said about viscosity as a parameter having basically important significance for small-scale motions, we can conclude what value of the Reynolds number is the lower limit of development of turbulence. With respect to the upper limit, it must be remembered that an important parameter in this case is the energy flux itself, as already mentioned. Of still greater importance is the applicability of the Reynolds criterion to a compressible liquid, i.e., to a gas. Theory and experience show that the criterion also is applicable in this case, but the compressibility factor must be taken into account.

In astrophysics (and geophysics) it is necessary to deal with two possible forms of compressibility: stratified and gasdynamic. The first is associated with thick layers of gas and the gravity effect; as a result its density becomes a function of the vertical coordinate. The second is caused by the velocity value. When the gasdynamic approaches the velocity of sound, and especially when it exceeds it, that is, when the Mach number Ma = v/c_s (c_s is the speed of sound) attains appreciable values, compressibility becomes a completely real value, and it must be taken into account.

In the interior of the sun, we must more often deal with stratified compressibility. Many investigators (Refs. 16, 11) have pointed out the limited applicability of the Reynolds criterion under these conditions and the preferability of the Richardson criterion

$$\mathrm{Ri} = \frac{g}{T} \cdot \frac{\left(\frac{\partial T}{\partial z} + \gamma_a\right)}{\left(\frac{\partial v}{\partial z}\right)^2} < 1, \tag{2.4}$$

where γ_a is the adiabatic vertical temperature gradient, g is the acceleration of gravity, v is horizontal velocity. The sense of this criterion is that the flux will remain turbulent if the transfer of energy from the higher terms of the turbulent hierarchy to the lower terms is not less than the work which should be performed against the force of gravity in order to maintain the turbulent element in the atmosphere. Unity on the right-hand side is obtained as the ratio of the vertical and horizontal diffusions when there is a rank coincidence of their values. It is therefore clear that the criterion is quite suitable for a medium with stratified density. With respect to the significance of the Richardson criterion, we can agree both with D. L. Laykhtman, who feels that this criterion characterizes the upper limit of turbulence to the same degree as the Reynolds criterion determines its lower limit, and with L. S. Gandin, who feels that the Ri number is an internal parameter differing appreciably from other similarity criteria (Ref. 15).

The introduction of the concept of the Reynolds number makes possible the estimation of the value of the coefficients of kinematic and dynamic turbulent viscosities (the latter is better known by the term "exchange coefficient"). In actuality, the Reynolds number is a dimensionless value, in whose denominator we have kinematic viscosity. The dimensionality of the numerator, therefore, also is the dimensionality of kinematic viscosity, which of course is directly obvious. Thus, the numerator in the Reynolds criterion also expresses viscosity. It therefore can be concluded that the expression standing in the numerator of the Reynolds criterion is turbulent kinematic viscosity, usually denoted K.

Thus,

$$K = \nu \mathrm{Re}. \tag{2.5}$$

K multiplied by ρ gives the exchange coefficient:

$$A = K\rho. \tag{2.6}$$

Whereas the turbulent viscosity is dependent only upon the velocity and scale of the pulsations, the exchange coefficient also is dependent upon density. In the subphotospheric layers it will therefore be different at different depths, even with the same scale and an identical order of velocities. For example, in a hydrodynamic study of phenomena of the scale of sunspots, the characteristic dimensions are of the order of 10^9 cm. Therefore, even at a velocity of the order of 1 m/sec = 10^2 cm/sec, which apparently is much too low, the coefficient of turbulent viscosity (we will omit the word "kinematic") will be of the order of 10^{11} $cm^2 \cdot sec^{-1}$.

In the third section of this chapter we discuss general circulation on the sun. The characteristic scale of this phenomenon is approximately $5 \cdot 10^{10}$ cm, and the characteristic velocity is 10^2 $cm \cdot sec^{-1}$. Therefore, the coefficient of turbulent viscosity will be about 10^{13} $cm^2 \cdot sec^{-1}$ (to be more exact, in this case, $5 \cdot 10^{12}$ $cm^2 \cdot sec^{-1}$). It is this value which is used by investigators of the general circulation of the sun (Refs. 17-19).

It is true that Vasyutinskiy (Ref. 11) and Cowling (Ref. 8) dispute the possibility of a transition from gas kinetic to turbulent viscosity for motions commensurable with general circulation on the sun by use of the Reynolds number. They feel that whereas turbulent motion "should be more or less random", meridional circulation possesses a considerable degree of axial symmetry. However, the principal flux itself should in scale and velocity represent an example of turbulent flow, decaying into a series of large pulsations. The carrying off of these and smaller pulsations by the principal flux is no obstacle, as already pointed out, to the application of the theory of stationary turbulence.

The coefficients of kinematic turbulent viscosity, therefore, are of the order of 10^{12}-10^{13} for motions commensurable in scale and velocity with the characteristics of general circulation on the sun; these should be quite close to reality. Certain authors have obtained

different values: for example, Csada, who first obtained $K = 10^{12}$-10^{13}, later indicated a value of 10^{15}, which is difficult to substantiate while remaining within the framework of hydrodynamics (Ref. 20). On the other hand, Schoenberg (Ref. 21) cited $K = 10^{6}$-10^{7} for a turbulent element, in order of magnitude equal to a granule (10^{8} cm), and for a rather considerable velocity, which is very strange. As already noted, the exchange coefficient obviously is dependent on the depth at which the turbulent exchange occurs. For example, if the coefficient of turbulent viscosity is equal to 10^{12}, at a depth of 1,000 km $A = 1.58 \cdot 10^{6}$, at a depth of 10,000 km $A = 7.59 \cdot 10^{6}$, and at a depth of 100,000 km $A = 1.78 \cdot 10^{9}$. All of these values exceed by many orders of magnitude the coefficients of gas kinetic viscosity.

We will now discuss turbulent heat conductivity. It is necessary to distinguish vertical and horizontal heat conductivity, or to express it differently, vertical and horizontal exchange. We will consider the problem of vertical exchange later, since it is associated closely with the convection problem. Horizontal turbulent heat exchange, or macroturbulent heat exchange, as it usually is called, can be computed in the following manner:

$$Q = -c_p K \rho \frac{1}{r} \frac{\partial T}{\partial \varphi}. \tag{2.7}$$

This formula obviously gives the heat flux transported by the macroturbulent method. The term $\frac{1}{r} \frac{\partial T}{\partial \varphi}$ represents the meridional temperature gradient. The heat capacity in the case of a constant pressure c_p in the subphotospheric layers changes very greatly with depth, attaining maximum values where there is the most intense ionization of hydrogen. We will give attention to this problem when discussing subphotospheric convection. We will now use for c_p a value close to that which corresponds to conditions of complete ionization of hydrogen at a sufficiently high temperature. For this case it is possible to assume $c_p = 5$ cal/degree, assuming further that the meridional temperature gradient is $2.62 \cdot 10^{-12}$ degree/cm (Ref. 22), and applying it to a depth

$z \approx 5 \cdot 10^8$ cm, where hydrogen can be considered virtually entirely ionized. Since there $\rho = 4.47 \cdot 10^{-6}$ g·cm^{-3}, when $K \approx 5 \cdot 10^{12}$, we obtain $|Q| = 1.26 \cdot 10^4$ ergs·cm^{-2}·sec^{-1}.

In discussing the electromagnetic properties of the subphotospheric gas it is necessary to mention the subphotospheric magnetic fields. This problem can be approached both deductively and inductively. In the inductive approach it is necessary to represent the structure, strength, etc., of the fields known in the photosphere as they are in the subphotospheric layers, for example, on the basis of admissible hypotheses it is necessary to represent the characteristics of the fields of spots, weak fields, etc., in the subphotospheric layers.

In the deductive approach it is necessary to proceed from more general considerations. The structure of magnetic fields of sunspots in layers accessible to observation is known rather well. At the center of the core the field is vertical, but with approach to its periphery the lines of force become increasingly inclined to the sun's surface. At the inner boundary of the penumbra the angle of inclination is about 50°. As is well known, the strength of the fields of spots falls in the range 100-4000 oe (extreme values); it is maximum at the center of the spot and decreases toward the boundary in approximate agreement with the Broxson formula (Ref. 22):

$$H_r = H_0\left(1 - \frac{r^2}{b^2}\right), \tag{2.8}$$

where b is the radius of the spot, and H_0 is the strength at the axis.

From the change of the recorded strength of the magnetic field with movement of the spots from the center to the limb of the disk, and also by other methods, it was possible to determine the value of the vertical gradient of the magnetic field. This value as obtained from observations by different authors is somewhat different, from 0.5 oe/km (Nicholson, Ref. 23) to 5.7 oe/km (Hautgast and Sluiters Ref. 24). Mattig, who also investigated the magnetic fields in spots theoretically in Ref. 25, concludes that the vertical gradient of magnetic field strength in a sunspot can be determined from the formula:

$$\frac{dH}{dz} = \frac{4H_0}{b_0}, \tag{2.9}$$

where H_0 and b_0 are the values of the field strength and the radius of the spot at the level of formation of that spectral line from which

the magnetic field was determined, that is, in the relatively high layers. Formula (2.9) gives 0.8 oe/km for the gradient, which agrees rather closely with the results obtained by Mattig (0.5-1 oe/km). However, it is risky to extrapolate such a value of the gradient to the subphotospheric layers. Unfortunately, at the present neither theory nor, of course, observations, can provide anything more satisfactory.

In a recent study by R. N. Ikhsanov (Ref. 26), devoted to the problem of the magnetic energy of sunspots, the gradient of the magnetic field with depth in the subphotospheric layers is therefore assumed equal to its photospheric value (0.5 oe/km). Ikhsanov concludes that the magnetic field of a spot intensifies to a certain depth, z_0, dependent upon the area of the spot; for the smallest spots it is $6 \cdot 10^7$, and for the largest it is $3 \cdot 10^9$ cm. The field then is propagated in depth, attaining a depth equal to $2z_0$, that is, from 10^8 to $6 \cdot 10^9$ cm (the approximate depth of the principal zone of convection). This is what can be said of the magnetic fields of spots. The subphotospheric structure of bipolar magnetic regions (invisible spots) is unknown, but it can be assumed that it has much in common with the subphotospheric structure of the fields of the spots themselves. The magnetic field also should increase with depth, and this increase can be more significant. The magnetic field of such "invisible spots" is of the order of 200 oe (Ref. 27). It is known that these fields frequently appear at the site of earlier existing spots or where spots soon will appear. According to Cowling (Ref. 28), the magnetic fields of spots are destroyed only after a very long time and they can be seen to disappear to some depth beneath the photosphere. Provided that this depth is not considered too great (if it was very great the field strength at the level of the spectral lines from which strength is determined would be less than 200 oe), and assuming that the field there retains its initial value, or one close to it, it is found that the vertical gradient of the magnetic field in the "invisible" spot will be nearly as great as in the visible spot.

The problem of the change of the strength of weak unipolar fields with depth still has not been investigated, and it is difficult to say anything about the character of such fields in the subphotospheric layers. S. B. Pikel'ner, in Ref. 10, notes correctly that "a weak field can develop in the most turbulent medium". Such fields apparently can be formed in the subphotospheric layers rather frequently. The so-called "general magnetic field of the sun" is traced in the polar regions, where its strength is about 1 oe. With respect to this field, a considerable number of mutually exclusive conclusions have been drawn and many hypotheses advanced.

At the present time most investigators are inclined to feel that this is a weak dipole field. The determination that the sun's general field is variable was an important observational fact; from 1952 through 1957 the north magnetic pole was situated in the northern hemisphere, but in 1957 there was a change, and the north magnetic pole moved into the southern hemisphere, where the south magnetic pole had been situated; the south magnetic pole appeared in the northern hemisphere a year later. The change, therefore, occurred in the epoch of the maximum of the 11-year cycle, whereas the change of sign of the leading spot in bipolar groups of a particular hemisphere occurs at the time of the minimum with a transition from cycle to cycle. The strength of this field is not constant and the change of its sign, in Pikel'ner's opinion, is evidence that this field has a superficial character, since quite like as occurs in magnetic stars, the variability of the field can be attributed to the circulation of gas. How deeply can such a field penetrate into the sub-photospheric layers is difficult to say. It is only obvious that this problem is associated closely with the problem of the thickness of the layer of the sun in which sufficiently strong circulation occurs.

That which has been stated so briefly exhausts what is known from observations, and we will now discuss the problem of what deduction can tell us about regular fields. It would be best to begin here with Alfven's concept, although it now has become outdated (Ref. 29). Alfvén assumed that the general magnetic field of the sun at a great depth and in the peripheral layers is different in its structure. It is uniform in the deep layers and is a dipole field in the layers closer to the surface. A more modern concept is one developed for the most part by Elsasser, but also by Bullard and Parker (Refs. 30-32).

It is assumed that the principal form of a general magnetic field is a field of the dipole type, called poloidal; in this case, the axis of the field coincides with the axis of rotation of a celestial body and the lines of force lie in the meridian planes of this body. If the rotation of the body was rigid, the field would retain its form and orientation. But for the rotation of the stars, including the sun, it is characteristic that angular velocity is dependent both on the distance to the center and on astrographic latitude. Even the first dependence leads to curvature of the lines of force of the poloidal field and their subsequent spiraling. Such a spiral line of force can be broken down into the line of the initial poloidal field, and a line having the form of a circle with its center at the axis of rotation. Therefore, there will be two families of lines: of a poloidal field, and of a second field having the form of a torus and called toroidal. The toroidal field, therefore, is induced; it is far stronger than the initial poloidal field. The toroidal field develops at great depths, where the difference between the angular velocities and their values at the surface attains high values. It will be demonstrated below that the difference

in the angular velocities of the deep and surface layers is less than assumed previously. This somewhat complicates the formation of toroidal fields.

Toroidal fields emerge at the surface in those cases when something conveys them there. The most effective mechanism of such a transfer is convection. It must be taken into account that toroidal fields possess buoyancy, since the gas pressure within the torus is less than around it (the pressure excess balancing it with the external pressure is created by magnetic pressure), and when the temperature is equal to that of the surrounding medium the density of the torus will be less, which ensures the appearance of the Archimedean force. According to Parker, the magnetic fields of spots, and also other magnetic fields appearing on the sun's surface, in fact, constitute the results of emergence of toroidal fields in the surface layers (Ref. 32).

Thus, whatever approach is used to the problem of solar magnetic fields, either induction or deduction, there is no difference and there can be only a single conclusion: the subphotospheric layers, to use figurative speech, are saturated with various magnetic fields.

Section 2. Convection in the Subphotospheric Layers and Present-Day Concepts Concerning Subphotospheric Stratification

Principles of Convection Theory

We will now proceed directly to the problem of convection. Since this phenomenon plays a decisive role in processes transpiring in a layer of more or less significant thickness, adjacent to the photosphere from below, we will discuss it in greater detail. Convection, or to be more exact, free thermal convection, is one of the forms of gravitational instability. If the layer of liquid or gas is heated from below and heat conductivity cannot ensure the equalizing of the heat, this equilization is accomplished by vertical macromotions that constitute convection.

We will first consider convection in an ideal gas, and will assume that no heat conductivity mechanisms operate, but that only motion generated by the difference of temperatures, that is, convection itself, equalizes this temperature difference. Since the concept of the vertical adiabatic temperature gradient is of fundamental significance in convection theory, we will show the means by which the corresponding expression is derived. For the adiabatic process, the equation representing the first law of thermodynamics obviously has the following form:

$$c_p dT - R\frac{T\,dp}{p} = 0. \qquad (2.10)$$

Here R is the specific gas constant, equal to $\mathfrak{R}/\mu$, where $\mathfrak{R} = 8.31 \cdot 10^7$ (c_p is expressed in mechanical units, therefore, there is no thermal equivalent of the work usually entering as the coefficient of the second term).

After dividing both sides of (2.10) by $c_p T dz$, we have

$$\frac{dT}{T\,dz} = \frac{R}{c_p}\frac{dp}{p\,dz}. \tag{2.11}$$

We now take the equation of hydrostatic equilibrium

$$dp = -g\rho dz \tag{2.12}$$

and the equation of state, which we apply to the atmosphere, having the temperature T', surrounding our mass of gas

$$\rho = \frac{p}{RT'}. \tag{2.13}$$

Substituting (2.12) into (2.13), we obtain

$$\frac{dp}{dz} = -\frac{pg}{RT'}. \tag{2.14}$$

The pressure in the atmosphere surrounding the gas can be considered identical. Substituting (2.14) into (2.11), we obtain

$$\frac{dT}{dz} = -\frac{g}{c_p}\frac{T}{T'}. \tag{2.15}$$

Assuming the temperature differences between the mass of gas and the surrounding atmosphere to be insignificant, we obtain

$$\frac{dT}{dz} = -\frac{g}{c_p}. \tag{2.16}$$

The modulus of the expression on the right-hand side of (2.16) also is the value of the adiabatic temperature gradient. The sign obviously indicates that temperature decreases with height.

We will now consider the conditions for vertical stability of the gas. We will assume that certain mechanical forces have uplifted a certain mass of gas. The process of expansion of this mass of gas can be considered adiabatic, that is, the temperature decrease in this mass of gas will occur in accordance with the adiabatic gradient.

We will assume that in the medium (also gaseous) surrounding the considered mass of gas, the temperature decrease will occur in accordance with a gradient, which is less than adiabatic. This means that the temperature in the medium decreases with height more slowly than would be the case if the decrease occurred in accordance with the adiabatic gradient. Then, at a certain height, the mass of gas will be colder than the medium surrounding it, and since the pressure in the mass of gas and in the surrounding medium are the same, it is easy to find from the equation of state that

$$\frac{\rho_1}{\rho_2}=\frac{T_2}{T_1}, \qquad (2.17)$$

where ρ_1 and T_1 apply to the mass of gas, and ρ_2 and T_2 apply to the surrounding medium. In addition, as we have established, $\rho_1 > \rho_2$, that is, the density of the ascending mass being the colder, will be greater than the density of the surrounding medium. But then this mass obviously should "sink" into the surrounding medium, that is, return to that level from which it began its ascent.

There will be a completely similar situation in a case when the mass of gas descends under the influence of mechanical forces, it being compressed in accordance with the adiabatic law. In this case it will be heated more intensely than the surrounding medium, will become less dense, and will tend to return to its initial position, i.e., in this particular case, rise upward. Thus, if the vertical temperature gradient in a particular medium is less than the adiabatic gradient, a particle of gas experiencing vertical movement under the influence of mechanical forces will tend to return to its former position. Such a stratification is called stable. Thus, if $\frac{dT}{dz}<\Gamma_a$, there is stable equilibrium (Γ_a is the adiabatic gradient).

We will now assume that the vertical temperature gradient of the medium is precisely equal to the adiabatic temperature gradient. Then the rising or descending mass of gas, moving under the influence of mechanical factors, will have the same density at all levels as the surrounding medium. This means that this mass of gas will be held at all levels as soon as the influence of the mechanical force imparting some vertical motion to this mass is removed.

In this case there obviously will be neutral equilibrium, that is, if $\frac{dT}{dz} = \Gamma_a$ there will be neutral equilibrium.

Finally, we will assume that the vertical temperature gradient of the medium is greater than adiabatic. Then the rising mass is cooled less than the surrounding medium, the density of this mass becomes less than the density of the medium, and the mass of gas begins to become buoyant. On the other hand, if some mass of gas descends in a medium under the influence of a mechanical force, the vertical temperature gradient of the medium being greater than the adiabatic gradient, this mass of gas will be heated more slowly than the surrounding medium, and will fall downward. The equilibrium in this case obviously will be unstable, which can be expressed as follows: if $\frac{Td}{dz} > \Gamma_a$, the equilibrium is unstable.

This case is of primary interest because it is in fact the case of free thermal convection. Thus, free convection is the term given to the vertical motions of gas (or liquid) masses occurring under conditions of unstable equilibrium, i.e., when $\frac{dT}{dz} > \Gamma_a$.

We will now consider a case in which the mass of gas does not receive a mechanical impulse directed upward or downward, but in which its temperature at the initial time differs from the temperature of the surrounding medium. We will find what vertical acceleration $\frac{d^2z}{dt^2}$ (t is time) is equal to in this case. Assume that the weight of an individual gas particle is F_i, the volume occupied by the particle is v, the density is ρ_i, and the temperature T_i. Assume the weight of the surrounding gas in this same volume, v, to be F_a, its density to be ρ_a, and its temperature to be T_a. The weight of the gas particle in the medium surrounding it obviously will be

$$F_i - F_a.$$

When $F_1 < F_2$ there will appear an upward-directed lifting force, F; thus,

$$F = F_a - F_i$$

(direction upward is considered positive). Replacing F_a and F_i by their

values in accordance with the expressions

$$F_a = gv\rho_a; \quad F_i = gv\rho_i,$$

we obtain

$$F = gv(\rho_a - \rho_i). \tag{2.18}$$

But it is obvious that

$$F = \rho_i v \frac{d^2z}{dt^2}, \tag{2.19}$$

$$\frac{d^2z}{dt^2} = g\frac{\rho_a - \rho_i}{\rho_i}, \tag{2.20}$$

or, using the equation of state,

$$\frac{d^2z}{dt^2} = g\frac{T_i - T_a}{T_a}. \tag{2.21}$$

Formulas (2.20) and (2.21) are called the Exner formulas. They show that the acceleration imparted to a particle of gas by convection is dependent on the difference in densities or temperatures of a particle and the surrounding gaseous medium. Depending on the sign of this difference the particle will move either upward (if $T_i > T_a$), or downward (if $T_i < T_a$). Since the change of temperature with height in a particle of gas can be considered adiabatic, and in the surrounding medium it can occur with any vertical temperature gradient, then:

$$T_i = T_{i0} - \Gamma_a z, \tag{2.22}$$

$$T_a = T_{a0} - \Gamma z, \tag{2.23}$$

where T_{i0} and T_{a0} are the temperatures at the initial level.

After substituting (2.22) and (2.23) into (2.21), we obtain

$$\frac{d^2z}{dt^2} = g\frac{T_{i0} - T_{a0}}{T_a} + g\frac{\Gamma - \Gamma_a}{T_a}z. \tag{2.24}$$

Thus, the acceleration of convection is dependent on two factors: the initial difference of temperatures and stratification.

The Schwarzschild criterion is the specific form of the criterion of convective instability for astrophysics. The adiabatic gradient and its radiative vertical gradient can be compared. By denoting the second

$\left|\frac{dT}{dz}\right|_{rad}$ and the first $\left|\frac{dT}{dz}\right|_{ad}$ it is possible to write the condition for

the existence of stable convective currents in the form

$$\left|\frac{dT}{dz}\right|_{ad} < \left|\frac{dT}{dz}\right|_{rad}. \tag{2.25}$$

This is one of the forms of the Schwarzschild criterion (Ref. 33). Usually it is given a different form, that is, ρ is excluded from the equation of state and the equation of hydrostatic equilibrium, giving the differential form of the barometric formula

$$\frac{dp}{\rho dz} = -\frac{g\mu}{RT} \tag{2.26}$$

(μ is the molecular weight).

Multiplying both sides of this expression by dT, we obtain

$$\frac{dT}{dz} d\ln p = -\frac{g\mu}{R} d\ln T. \tag{2.27}$$

Substituting this into the expression for the Schwartzschild criterion, we obtain

$$\left|\frac{d\ln T}{d\ln p}\right|_{ad} < \left|\frac{d\ln T}{d\ln p}\right|_{rad}. \tag{2.28}$$

In certain cases this formula is more convenient for astrophysical purposes. In particular, it makes it possible to represent the adiabatic gradient through the ratio of heat capacities at a constant pressure and constant volume $c_p/c_v = \gamma$. From the equation for the adiabatic process

$$T = \text{const}\, p^{\frac{\gamma - 1}{\gamma}}, \tag{2.29}$$

after being put into logarithmic form and being differentiated, we obtain

$$\left|\frac{d\ln T}{d\ln p}\right|_{ad} = \frac{\gamma - 1}{\gamma}. \tag{2.30}$$

This is the form of the expression for the adiabatic gradient usually used in astrophysics. Since for a monatomic gas $\gamma = 5/3$, $\frac{\gamma-1}{\gamma} = 0.40$. However, this value does not remain constant in the subphotospheric layers because it is strongly dependent on ionization.

We will now consider the development and persistence of convection in a real gas, recalling that viscosity and heat conductivity must be taken into account. Viscosity will be a resistance to convective flow; therefore, if convective flow developed in the absence of viscosity, and with a very insignificant excess of the temperature gradient above the adiabatic gradient, this excess should be more considerable in the presence of viscosity. Heat conductivity will tend to equalize the temperature difference; therefore, the most effective mechanism will be that which makes the development of convection most difficult. In astrophysics the most important condition is the transfer of heat by radiation, i.e., radiative heat conductivity.

Rayleigh (Ref. 34) formulated the problem of finding a dimensionless number that would characterize a particular medium, in the sense of conditions for the development of convection, the same as the Reynolds number characterizes a medium in the sense of conditions for the development of turbulence. Thus, there should be not only such an expression, but also its numerical value, which would play the role of a critical value, so that if in a particular liquid (gas) this threshold is exceeded, convection will develop. The expression for this dimensionless number is found by investigating a certain volume of gas in a gaseous state. With ascent it experiences an adiabatic change of temperature, at the same time that the temperature in the medium itself changes with height, in accordance with some vertical gradient $\frac{dT}{dz}$. Then with ascent, by a unit length, the difference in temperatures of the considered volume of gas and the medium changes by the value

$$\left[\left(\frac{dT}{dz}\right)_{ad}-\left(\frac{dT}{dz}\right)\right].$$

If the ascent of an element of the volume of gas occurs to the height l, the difference of temperature between this element of volume and the medium will be

$$l\left[\left(\frac{dT}{dz}\right)_{ad}-\left(\frac{dT}{dz}\right)\right]=l\Delta\,\mathrm{grad}\,T. \tag{2.31}$$

The coefficient of expansion of the medium then is introduced into the problem. At constant pressure, that is, when nothing hinders expansion, the coefficient obviously will be

$$\alpha=\frac{1}{v}\left(\frac{dv}{dT}\right)_p. \tag{2.32}$$

Here v is the specific volume. From the Exner formula (2.20), which is used in the form

$$\frac{d^2z}{dt^2} = g\frac{\bar{\rho} - \rho}{\rho},$$

where $\bar{\rho}$ is the density of the medium, ρ is the density of an element of the volume, by replacing densities with specific volumes it is possible to obtain the following expression:

$$\frac{d^2z}{dt^2} = \frac{v - \bar{v}}{\bar{v}}.$$

But it is obvious that

$$v = \bar{v}(1 + \alpha l\Delta \operatorname{grad} T),$$

hence

$$v - \bar{v} = \bar{v}\alpha l\Delta \operatorname{grad} T.$$

Thus, the force per unit volume will be

$$F = \rho lg\alpha\Delta \operatorname{grad} T. \tag{2.33}$$

As follows from the Navier-Stokes equation, the viscosity has the form

$$|K_r| = \eta\nabla^2 u,$$

where u is velocity, η is the coefficient of dynamic viscosity. In order of magnitude this is equal to

$$O(K_r) = \frac{\eta u}{l^2}, \tag{2.34}$$

or, converting to kinematic viscosity,

$$O(K_r) = \frac{\nu\rho u}{l^2}. \tag{2.35}$$

At the threshold of development of convection the lifting force should be precisely equal to viscosity (internal friction). At this time the motion occurring under the influence of the mechanical impulse is inertial (since the acceleration imparted by convection at this time is equal to zero). Consequently, equating (2.33) to (2.34), and thereby determining velocity, we have:

$$u = \frac{\alpha g l^3\Delta \operatorname{grad} T}{\nu}. \tag{2.36}$$

The influence of heat conductivity is essentially as follows: convection should equalize the temperature difference $l\Delta \operatorname{grad} T$ during the time of passage of an elementary volume along the path l.

It is obvious that at the velocity u, this time has the order of magnitude l/u. Another possible mechanism of equalization of temperature, as previously mentioned, is heat conductivity. Regardless of what heat

conductivity is involved--gas kinetic or radiative--the rate of temperature equalization will be determined directly by the value called thermal conductivity, having the form

$$a = \frac{\lambda}{c_v \rho} \tag{2.37}$$

(λ is heat conductivity).

This value has the dimensionality of kinematic turbulent viscosity, which is obvious from the relation $\lambda = 5/3\ c_v \eta$, correct for a monatomic gas. Consequently, $[a] = \mathrm{cm}^2 \cdot \mathrm{sec}^{-1}$. The order of magnitude of the time, determined from the heat conductivity equation or found from the above-mentioned dimensionality, will be

$$t = \frac{l^2}{a}.$$

By comparing the times of passage of the path l involved in convection we obtain a ratio of considerable importance for convection. In actuality, by denoting the first time, t, and the second, t_1, we see that for development of convection it is necessary to satisfy the condition $t << t$. The ratio t_1 to t obviously is equal to:

$$\frac{l^2}{a} : \frac{l}{u} = \frac{lu}{a}. \tag{2.38}$$

By replacing u by its value from (2.36), we obtain the dimensionless Rayleigh number

$$R = \frac{\alpha g l^4 \Delta \operatorname{grad} T}{\nu a}. \tag{2.39}$$

The critical value of the Rayleigh number R_c corresponds to the threshold of convection. Thus, convection is stable when R R_c. The mean value of R_c is assumed to be 1,700; however, its more precise value is dependent on the boundary conditions. The case R_c = 1,700 corresponds to the presence in the liquid or gas of two limiting rigid surfaces. When there is one free boundary, R_c = 1,101. In the absence of boundaries, or to be more exact, when there are two free boundaries, R_c = 658, as shown by the theory developed by Rayleigh. When R only slightly exceeds R_c, convection will be stationary, i.e., the transfer of heat by convective currents will have the same order of magnitude as

the heat conductivity, and the lifting force will be of the same order as the viscosity. In this case a liquid will form a cellular structure, as shown by experiments performed for the first time by Benard (Ref. 35). The cells are hexagonal. The liquid rises at the center of each cell and descends along its sides. The phenomenon of the Benard cells is observed rather frequently under natural conditions: certain cloud genera, zones at sea, etc. Until recently it was assumed that the photospheric granules are an astrophysical example of the Benard cells, but in recent years this has been subjected to doubt (Ref. 36). Vasyutinskiy even is inclined to relate the Benard cells to certain geomorphological phenomena (Ref. 11). The most important parameter of the cells is the ratio of a side of the base to the height. There is a rather rigorous relationship here. When there are two rigid boundaries, the ratio is 1.34:1, when there is one rigid boundary, the ratio is 1.56:1, and finally, when there are two free boundaries, it is 1.89:1 (Ref. 10). The theory of this problem was considered not only by Rayleigh, but by Jeffreys, Brent, and others, as well. The others give a somewhat higher value of this ratio (about 3:1).

In analyzing this phenomenon, the theoreticians proceed on the basis of conditions under which convective motion is stable. In most cases the method of small perturbations is used, which inevitably involves appreciable simplifications of the problem. For example, Rayleigh assumed that the cells are tetrahedral, whereas actually they are hexagonal. It should cause no surprise that the theory of this complex phenomenon is currently not sufficiently developed to explain all of the experimentally observed characteristic peculiarities. If R is considerably greater than R_c, the form of the cells will be distorted sharply,

and with a further increase of R, convection, in general, will become disordered. In this case we refer to it as turbulent convection. As shown by the very form of the expression for the Rayleigh criterion, this value is particularly sensitive to an increase of l. It is easy to see that l corresponds to what is called the mixing length in turbulence theory. For liquids which are compressible in the stratified sense, that is, when there is a considerable drop in density in the layer, the mixing length is of the same order of magnitude as the height of the homogeneous atmosphere. It can be either equal to it or exceed it by a factor of 1-1/2 to 2. The relationship between the Rayleigh and Reynolds numbers can be obtained by a comparison of expressions (2.39) and (2.3). It therefore follows that

$$R = \frac{ul}{a}. \tag{2.40}$$

Multiplying and dividing by the kinematic viscosity, and substituting expression (2.3) for the Reynolds number, we obtain

$$R = \frac{\mathrm{Re}\,\nu}{a}. \qquad (2.41)$$

Since ν and a have the same order of magnitude, in theory we should have $R \simeq \mathrm{Re}$, but, in actuality, experience shows that the following relation should be closer to the truth

$$R = 100\ \mathrm{Re}. \qquad (2.42)$$

It therefore follows that turbulent convection should arise when $R > 100{,}000$, since turbulent motion arises when $\mathrm{Re} > 1{,}000$. However, in many cases nonstationary convection arises when $R > 50{,}000$.

An important case in certain astrophysical problems is that in which convection arises under conditions of already existing turbulence caused by large-scale motion at a considerable velocity. In this case, ν is replaced by the coefficient of kinematic turbulent viscosity, K. A similar value obviously should be introduced in place of a. After multiplying the numerator and denominator of (2.38) by ρ^2, and taking into account expression (2.6) for the coefficient of turbulent exchange, we obtain

$$R = \frac{\alpha g \rho^2 l^4 \Delta\ \mathrm{grad}\ T}{A^2}. \qquad (2.43)$$

Two expressions can be derived for the velocity of convection. These expressions are used in the theory of a convective zone.

The first expression is applicable to an ideal gas, that is, the expression makes no allowance for viscosity or heat conductivity. The lifting force (Archimedean force) obviously is equal to the difference in weight between the gas involved and the displaced gas. When there is an equality of volumes we can write:

m_1 = the volume ρ_1, m_1 and m_2 are masses,

m_2 = the volume ρ_2,

or, converting to weights:

$m_1 g$ = g volume ρ_1,

$m_2 g$ = g volume ρ_2.

Taking the lifting force with a minus sign, we have:

lifting force = -g volume $\Delta\rho$. (2.44)

After the equation of state, $p = \rho RT$, is put into logarithmic form and differentiated, we have ($R = \frac{\Re}{\mu}$):

$$\frac{\Delta\rho}{\rho} = -\frac{\Delta T}{T}. \tag{2.45}$$

Now the lifting force can be expressed as follows:

$$\text{lifting force} = g \text{ volume } \rho\frac{\Delta T}{T} = mg\frac{\Delta T}{T}. \tag{2.46}$$

The sense of the volume ΔT is that it is the heating of an elementary volume in relation to the surrounding medium. If the temperature in an elementary volume is denoted T', and the temperature of the medium as T, it is possible to write:

$$dT = \left(\frac{\partial T'}{\partial z} - \frac{\partial T}{\partial z}\right) dz. \tag{2.47}$$

The finite difference of the gradients is expressed as follows:

$$\Delta T = \int_0^z \left(\frac{\partial T'}{\partial z} - \frac{\partial T}{\partial z}\right) dz. \tag{2.48}$$

In order to obtain the work performed by the lifting force, we multiply both sides of (2.46) by the height z; replacing ΔT by its value from (2.48), we obtain

$$\mathfrak{A} = mg\frac{\int_0^z \left(\frac{\partial T'}{\partial z} - \frac{\partial T}{\partial z}\right) dz}{T}. \tag{2.49}$$

Equating to this expression the value of the kinetic energy, $mu^2/2$, we have

$$u = \sqrt{\frac{2g}{T}}\sqrt{\int_0^z \left(\frac{\partial T'}{\partial z} - \frac{\partial T}{\partial z}\right) z\,dz}. \tag{2.50}$$

If

$$\left(\frac{\partial T'}{\partial z} - \frac{\partial T}{\partial z}\right) = \text{const (Ref. 37)},$$

then

$$u=\sqrt{g}\sqrt{\frac{\frac{\partial T'}{\partial z}-\frac{\partial T}{\partial z}}{T}}\cdot z, \tag{2.51}$$

that is, velocity is proportional to distance from the initial point. In inspecting the mixing length l, it is possible to use as mean velocity its value when z = l/2; hence

$$u=\sqrt{g}\sqrt{\frac{\frac{\partial T'}{\partial z}-\frac{\partial T}{\partial z}}{T}}\cdot\frac{l}{2}. \tag{2.52}$$

If free convection occurs under conditions of turbulent friction, the expression for its velocity will be somewhat different. We will now consider a turbulent element with the cross section s and the height l (height has the characteristic dimension of the mixing length). This corresponds to what was stated previously concerning the coincidence of the orders of magnitude of the mixing length, and the height of the homogeneous atmosphere in compressible media. In this case the lifting force $\mathfrak{F}$ is written as follows:

$$\mathfrak{F}=-slg\Delta\rho=-\rho sgl\,\frac{\Delta\rho}{\rho}. \tag{2.53}$$

It follows from formula (2.45) that $\frac{\Delta\rho}{\rho}=-\frac{\Delta T}{T}$; hence

$$\mathfrak{F}=s\rho gl\,\frac{\Delta T}{T}. \tag{2.54}$$

Under stationary conditions the lifting force should be equal to the force of the resistance, which in the case of turbulent friction can be assumed proportional to density, the area of the cross section and the square of velocity, that is

$$K_r=s\rho u^2. \tag{2.55}$$

Equating (2.54) to (2.55), we find that

$$u=\sqrt{gl\frac{\Delta T}{T}}. \tag{2.56}$$

Thus, in this case the characteristic length enters under the radical.

An important concept in convection theory, especially in connection with the theory of a convective zone, is the determination of the connective flow. It is expressed as follows:

$$\pi F_k=c_p\rho\Delta T u. \tag{2.57}$$

The dimensionality of πF_k is either $cal \cdot cm^{-2} \cdot sec^{-1}$, or $ergs \cdot cm^{2}$ $\cdot sec$, depending on what units are used to express c_p.

Convective power is another concept important in convection theory. We will now discuss the model shown in Figure 21. At the level denoted on the diagram by z, we will define a certain area, (z), and will consider the work performed at the time of vertical movements in the layer sdz. If we denote the temperature of the ascending and descending elements by T in the Exner formula (2.21), and the same in the surrounding medium by $\bar{T}$, then

$$\frac{d^2z}{dT^2} = g\,\frac{T - T}{T}.$$

Bearing in mind that the work is produced by a force computed per unit mass, it is possible to replace acceleration by the derivative of energy along the path

$$\frac{dE}{dz} = g\left(\frac{T}{T} - 1\right),$$

or

$$dE = g\left(\frac{T}{T} - 1\right)dz. \qquad (2.58)$$

We will denote by $m\uparrow$ a particle moving upward, and by $m\downarrow$, a particle moving downward. Since the first travels a path +dz, and the second -dz, the total work is represented as the algebraic sum of two kinds of work:

$$\delta E = g\,\frac{T_{\uparrow} - T}{T}\,m_{\uparrow}dz - g\,\frac{T_{\downarrow} - T}{T}\,m_{\downarrow}z. \qquad (2.59)$$

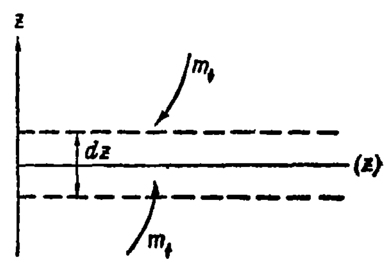


Figure 21. Diagram explaining the derivation of convective power.

Here $T\uparrow$ is the temperature of a particle moving upward, and $T\downarrow$ of a particle moving downward; in addition, the masses of the particles are introduced. If the mixing length is introduced, which in theory can be different for ascending and descending particles, it can be postulated that one of the particles is lost from the medium at the level $z - l\uparrow$, and the other at the level $z + l\downarrow$. At the initial levels the particles had the temperature of the surrounding medium; then

$$\left.\begin{aligned} T_{\uparrow} &= \bar{T}_{(z-l_{\uparrow})} - l_{\uparrow}\gamma_{\uparrow}, \\ T_{\downarrow} &= \bar{T}_{(z+l_{\downarrow})} + l_{\downarrow}\gamma_{\downarrow}, \end{aligned}\right\} \tag{2.60}$$

or since

$$\bar{T}_{(z-l_{\uparrow})} = \bar{T}(z) - l_{\uparrow}\frac{\partial \bar{T}}{\partial z},$$

$$\bar{T}_{(z+l_{\downarrow})} = \bar{T}(z) + l_{\downarrow}\frac{\partial \bar{T}}{\partial z},$$

$$\left.\begin{aligned} T_{\uparrow} &= \bar{T}(z) - l_{\uparrow}\left(\frac{\partial \bar{T}}{\partial z} + \gamma_{\uparrow}\right), \\ T_{\downarrow} &= \bar{T}(z) + l_{\downarrow}\left(\frac{\partial \bar{T}}{\partial z} + \gamma_{\downarrow}\right). \end{aligned}\right\} \tag{2.61}$$

In general, the gradients $\gamma\uparrow$ and $\gamma\downarrow$ may not be equal. Substituting (2.61) into (2.59), and summing for the area s and the time t, we obtain

$$\delta E = -g\frac{dz}{T}\left[\left(\frac{\partial \bar{T}}{\partial z} + \gamma_{\uparrow}\right)\sum_{s,t} m_{\uparrow}l_{\uparrow} + \left(\frac{\partial \bar{T}}{\partial z} + \gamma_{\downarrow}\right)\sum_{s,t} m_{\downarrow}l_{\downarrow}\right]. \tag{2.62}$$

By introducing the following notations

$$\frac{\sum_{s,t} m_{\uparrow}l_{\uparrow}}{st} = A_{\uparrow}, \tag{2.63}$$

$$\frac{\sum_{s,t} m_{\downarrow}l_{\downarrow}}{st} = A_{\downarrow}, \tag{2.64}$$

it is possible to write an expression for convective power in the layer (0, z) in the form

$$E = -g\int_0^z \left[\left(\frac{\partial \bar{T}}{\partial z} + \gamma_{\uparrow}\right)A_{\uparrow} + \left(\frac{\partial \bar{T}}{\partial z} + \gamma_{\downarrow}\right)A_{\downarrow}\right]\frac{dz}{T}. \tag{2.65}$$

Thus, convective power is the work done in a unit time during the vertical mixing of all masses participating in convective motion. The work usually is considered in a column with a unit cross section. The expression

$$A = A_{\uparrow} + A_{\downarrow} = \frac{1}{st}\left(\sum_{s,t} m_{\uparrow} l_{\uparrow} + \sum_{s,t} m_{\downarrow} l_{\downarrow}\right) \tag{2.66}$$

is called the exchange coefficient in the Schmidt form (Ref. 38). It has the same value as given by formula (2.6). It is obvious that m and l, characteristic for convection, can have different orders of magnitude for turbulence, especially for macroturbulent exchange. In this case, by expressing the exchange coefficient in Schmidt form, it is assumed that there is an equality of the "masses of eddies" moving in mutually opposite directions, and also equality of their dimensions. If it is assumed further that in steady-state exchange, one-half of all the eddies move in one direction, and the other one-half in the opposite direction, (2.66) for this case will be written as follows:[1]

$$A = \frac{\sum_{i} m_i l_i}{st}. \tag{2.67}$$

The value l in the first approximation is equal to the mean dimensions of the eddy characteristic for the transport of energy, but more rigorously it is equal to the mean distance covered by one eddy before its merging with another. This value sometimes is called the origin length.

An important form of heat transfer (and certain other characteristics) in geophysics is advection, or to be more precise, horizontal advection, since this is horizontal transfer, although it occurs in a layer having a definite vertical extent. This form of transfer can be of importance in astrophysics as well. If the geostrophic wind is used as a point of departure, that is, wind with inertial (uniform and linear) motion, the change of temperature with time will give the expression:

$$\frac{\partial T}{\partial t} = -\frac{1}{2\omega \cos\theta\rho}\frac{\partial p}{\partial n_1}\frac{\partial T}{\partial n_2}\sin\beta, \tag{2.68}$$

---

[1]This naturally does not mean that the convective current always is equal to zero; the eddies moving in one direction are of a greater capacity in relation to transported value than the eddies moving in the opposite direction and as always the flux is the difference between the two values.

where ω is the angular velocity of rotation of a celestial body; $\frac{\partial p}{\partial n_1}$ and $\frac{\partial T}{\partial n_2}$ are the pressure and temperature gradients, respectively, in a horizontal plane; θ is colatitude; and β is the angle between the pressure and temperature gradients. As shown in dynamic meteorology, formula (2.68) can be reduced by introducing the velocity of the geostrophic wind v_g, and the change of the angle between the pressure and temperature gradients with height to the form

$$\frac{\partial T}{\partial t}=\frac{2\omega\cos\theta}{g}v_g^2\frac{d\beta}{dz}. \tag{2.69}$$

It follows from formulas (2.68) and (2.69) that in the northern hemisphere, where $\omega\cos\theta > 0$ (θ changes from 0 at the north pole of the celestial body to π at the south pole), when $\beta > 0$, which corresponds to a leftward wind shear with height, there is an advection of cold, and when $\beta < 0$, there is an advection of heat. In this case there is a wind shear to the right. The value of the advection is proportional to the sine of the angle between the gradients, and also to the gradients themselves. All other conditions being equal, geostrophic advection will be maximal when the isotherms and isobars are mutually perpendicular. The dependence of the sign of change of temperature at the time of advection on the direction of wind shear with height is not entirely obvious from formula (2.68), without additional considerations, but we have no room to discuss them here (the reader should consult the cited literature on dynamic meteorology). This becomes more obvious from a consideration of formula (2.69), where it can be seen that the sign of advection is determined by the sign of $\frac{d\beta}{dz}$. From this same formula, it follows that there is a dependence of the advection value on the velocity of wind shear with height, and on the square of the value of the geostrophic wind. A more complete theory of heat advection in the atmosphere was developed by L. R. Rakipova (Ref. 39). In this theory the investigation was made for the general case of deviation of the temperature, wind velocity and heat flux values from their zonal values (that is, roughly speaking, from certain mean values for a particular latitude). A study is made of advection during a certain time in the entire column of gas (air, to be specific) and the change of the heat content in this column. The investigation is made in spherical coordinates, which is extremely important in connection with the role of this phenomenon in the general circulation of the atmosphere. But with respect to the details of the corresponding derivations, we again refer the reader to the original study (Ref. 39). However, we will cite the expressions derived by Rakipova for latitudinal (that is, zonal) and meridional advection components:

(1) latitudinal component

$$A_\lambda = c_p \left\{ -\frac{\alpha T}{g} \frac{\partial p'}{\partial \lambda} \Big|_z^0 - \frac{1}{R} \int_0^z \alpha \frac{\partial p}{\partial \lambda} dz + \frac{\alpha'}{g} \int_0^z T \frac{\partial p}{\partial \lambda} dz \right\}; \qquad (2.70)$$

(2) meridional component

$$A_\theta = c_p \left\{ \frac{\alpha'}{g} \int_0^z T \frac{\partial p}{\partial \lambda} dz + \frac{1}{2g\omega \cos\theta \sin\theta a_0^2} \left[\frac{T}{\rho} \frac{\partial p'}{\partial \lambda} \frac{\partial p'}{\partial \theta} \Big|_0^z - \right.\right.$$

$$\left.\left. - \int_0^z \frac{RT^2}{p} \frac{\partial p'}{\partial \theta} \frac{\partial^2 p'}{\partial z \partial \lambda} dz + 2\gamma \int_0^z \frac{1}{\rho} \frac{\partial p}{\partial \theta} \frac{\partial p}{\partial \lambda} dz \right] \right\}. \qquad (2.71)$$

Here p' is the deviation of pressure from the zonal value; λ is longitude; g is acceleration of gravity; R is the specific gas constant; α is the so-called circulation index, representing the ratio of the mean zonal velocity as a function of latitude to the distance of the point to the axis of rotation of the celestial body (radius of a parallel);

$-\alpha' = \frac{d\alpha}{dz}$, that is, the change of the index of circulation with height;

γ is the real vertical temperature gradient (stratification gradient); and, a_0 is the radius of the celestial body. It also is of interest to cite an expression for the change of the heat content of a column of the atmosphere with the height z during the time $t-t_0$. It is given by the expression

$$\Lambda = \frac{c_p}{t-t_0} \left\{ \frac{1}{g} (\bar{T} p' |_z - \bar{T} p' |_0) \Big|_{t_0}^t + \left(\frac{\gamma}{g} + \frac{1}{R} \right) \int_0^z p' \Big|_{t_0}^t \right\} dz. \qquad (2.72)$$

The horizontal advection of heat should differ appreciably from horizontal macroturbulent exchange. Roughly, the difference is that whereas in advection the transport of heat is along the lines of flow, in horizontal turbulent exchange the transport of heat is characteristically across the lines of flow.

We will now discuss the problem of forced convection. As repeatedly mentioned, free convection occurs only under conditions of unstable equilibrium, that is, when the stratification gradient of the medium is greater than adiabatic. Under such a situation it is sufficient to have an impulse directed upward for the mass of gas to begin to rise, or an impulse directed downward for a descending motion to begin.

As we have seen, in the place of a mechanical impulse there can be an initial difference between the temperatures of an element of gas and the medium surrounding it. We will now assume that the stratification is stable, that is, that the stratification gradient is less than the adiabatic gradient. But there is a certain mechanical factor overcoming the resistance of this stable medium and forcing the elements of the gas to rise upward and/or descend downward, that is, creating motions similar to convective. It is obvious that work is expended on this. This phenomenon is called forced convection. Forced convection can be caused not only by mechanical impulses, but by thermal factors as well.

As can be seen from formula (2.29), acceleration of convection is dependent on two factors: the initial difference of temperatures, and stratification. Under conditions of stable stratification, the convective acceleration imparted by the initial temperature difference will attenuate, but if this difference is maintained by a heat influx to the elements of gas, the convection will continue. Forced convection also has a certain significance in astrophysics, but its role should not be exaggerated, as was done several years ago by Schmeidler (Ref. 40), who postulated that the entire sun, with the exception of the photosphere, the subphotospheric zone of free convection, and the zones of free convection near the center of the sun (convective core), is in a state of forced convection whose energy source is the subphotospheric convective zone.

We will now consider the influence of external forces on convection. Bearing astrophysical problems in mind, it is necessary to discuss two forces: Coriolis force, associated with the rotation of a celestial body, and electromagnetic force. The influence of Coriolis force on convection already has been considered by Taylor (Ref. 41). He introduced a dimensionless criterion (Taylor number), having the form

$$T = \frac{4\Omega_z^2 d^4 \cos^2\theta}{\nu^2}, \tag{2.73}$$

where Ω_z is the vertical component of angular velocity of rotation of a celestial body. An increase of the value T leads to an increase of the value R_c, that is, delays the development of convection.

An investigation of this problem was made in 1947 in the USSR by L. S. Gandin (Ref. 42), and in 1953 in the United States by Chandrassekhar (Ref. 43). Gandin formulated the following problem: assume that above the surface of a celestial body, rotating with the angular velocity ω, there is a layer of gas of the fixed thickness, d, (for the sake of uniformity we use those notations which are used with the Taylor criterion, not those used by Gandin). The vertical temperature gradient in this layer of gas is denoted β. We will assume that β begins to

increase immediately in the entire layer. The problem is to find at which parallel convective motions begin initially, that is, in the case of a minimum value β. The investigation leads to the conclusion that the ratio of the horizontal dimensions of the developing convective circulations (cells) to the thickness of the layer is proportional to the cotangent of the latitude at which these circulations develop (or, accordingly, tan θ). The angular velocity of rotation of a celestial body, the thickness of the layer and kinematic velocity of the gas are related to astrographic latitude by the following relation

$$\mathrm{tg}^6\,\theta + \frac{3}{2}\,\mathrm{tg}^4\,\theta - \frac{1}{2} = \frac{2\omega^2 d^4}{\pi^4 \nu^2}. \tag{2.74}$$

As pointed out by S. B. Pikel'ner, the physical mechanism of influence of Coriolis force on convection is that under the influence of these forces there is a spiraling of horizontal flow in the convective cells, and as a result the entry of the gas to the center of the cell is attenuated, and it is necessary to have an additional increase for convection to develop. This effect, in view of the obvious dependence of the Coriolis force on latitude, will be related closely to the latter, which also is reflected in the Taylor criterion and in the expression on the right-hand side of the Gandin equation (2.74).

According to Chandrassekhar, the influence of T on convection becomes appreciable when $T > 10^3$, and when there are sufficiently high T values, it can be assumed that $R_{cr} \sim T^{2/3}$. In this case, the ratio of the horizontal dimension of the cells to the thickness d of the layer is proportional to $T^{1/6}$, i.e., the cells are deformed. The value of the Taylor number for convection on the sun applies for a layer thickness exceeding 100 km.

The problem of the influence of a magnetic field on convection is still more complex and has been less fully studied. Much attention has been given to this problem by Chandrassekhar (Ref. 44). In the case of an incompressible liquid, the critical value of the Rayleigh numbers R_{cr} is:

$$R_{cr} = \frac{(\pi^2 + a^2)\,[(\pi^2 + a^2)^2 + \pi^2 Q]}{a^2}, \tag{2.75}$$

where

$$Q = \frac{H_z^2}{4\pi\rho} \cdot \frac{d^2}{\nu\nu_m}, \tag{2.76}$$

and where H_z is the vertical component of the magnetic field, and it is this that is the obstacle to convection. The presence of a magnetic field increases the critical value of the Rayleigh number. Chandrassekhar gives a table showing the relationship of Q to R_{cr} for different boundary conditions (Table 19).

The presence of a field is reflected not only on the value R_{cr}, that is, on the threshold of convection, but also on the form of the convective cells. If it has vertical and horizontal components, the second component is no hindrance to convection, but facilitates the elongation of cells in a horizontal direction. If the field is vertical, with its intensification (or with an increase of Q for some other reason, such as increase of conductivity) the horizontal dimensions of the cells decrease. The electromagnetic factor then determines (all other conditions being equal) the transition of convection from stationary to nonstationary. Stationary convection occurs when $X < v_m$, where X is thermal conductivity, v_m is magnetic viscosity: $v_m \sim \frac{1}{\sigma}$. Since in astrophysics we deal with radiative thermal conductivity, $X \gg v_m$, which facilitates the development of nonstationary convection.

It must be remembered, however, that under definite conditions this nonstationary convection leads to the appearance of oscillatory instability. Actually, magnetic viscosity is determined by the conductivity of the medium, but field strength also is of appreciable importance; in the case of sufficiently large H (here H is the total value of the field), the field counteracts the displacement of the medium, leading to the appearance of Alfvén magnetohydrodynamic waves. Such oscillatory instability is manifested when

$$Q > \frac{27\pi^2 (\nu_m + \nu)}{4\nu}. \qquad (2.77)$$

The combination of a magnetic field in a conducting medium and rotation makes the convection exceptionally complex.

Principles of the Theory of the Convective Zone

1. Microscopic theory

In the course of the preceding discussion we have stated repeatedly that the development of convection in the subphotospheric layers is associated with the ionization of the gases present there, especially hydrogen, and secondarily, with helium. We will now discuss these problems in greater detail.

Table 19

Q	Two free boundaries	Two rigid boundaries	One free and one rigid boundary
0	657.5	1708	1101
100	2654	3757	1699
500	8579	10100	3586
1000	15210	17100	5613
10000	119800	124500	35040

The theory of the subphotospheric convective zone, or to be more precise, zones, can be broken down naturally into two parts: one part should consider atomic mechanisms leading to conditions under which a convective state develops, and the other, on the basis of the previously discussed general theory of convection, should lead to quantitative conclusions concerning subphotospheric conditions, and on the basis of these conclusions, yield the most up-to-date model of the structure of the subphotospheric layers, i.e., a model of the distribution of temperature, density, pressure, etc., with depth.

Turning to the Schwarzschild criterion, we see that the development of convection can be effected first by the increase of the radiative temperature gradient with depth, and, second, by a decrease of the value of the adiabatic gradient. It should be noted that the second means for the development of convection is purely astrophysical; under geophysical conditions the adiabatic gradient is essentially a completely constant value in that the heat capacity at a constant pressure remains unchanged within that gradiation of temperatures encountered under terrestrial conditions (the very high layers of the earth's atmosphere constitute a possible exception). Under astrophysical conditions a specific mechanism, such as ionization, is operative, leading to a change of the value of the adiabatic gradient.

We will first consider the behavior of the radiative temperature gradient. From the most elementary relation of photospheric theory, it follows that

$$T^4 = T_0^4\left(1 + \frac{3}{2}\tau\right). \qquad (2.78)$$

The temperature change with optical depth, therefore, will be

$$\frac{dT}{d\tau} = \frac{3T}{8\left(1 + \frac{3}{2}\tau\right)}. \qquad (2.79)$$

Since temperature and optical depth are on the right-hand side of this formula, the temperature change is dependent on the optical depth itself, and also on the coefficient of absorption, density and depth (since $d\tau = -\bar{\chi}\rho dz$).

One of the most complex problems is the change of the coefficient of absorption in the layers considered, relatively close to the photosphere. In the corresponding models of the sun, such computations nevertheless have been made. As an extremely rough approximation, the mean coefficient of absorption can be represented by the following interpolation formula:

$$\bar{\chi} = \bar{\chi}_0 \cdot \tau^m. \tag{2.80}$$

It is therefore possible to obtain:

$$\left(\frac{d \ln T}{d \ln p}\right)_{\text{rad}} = \frac{1}{4(1-m)} \cdot \frac{\tau}{\frac{3}{2}+\tau}. \tag{2.81}$$

In the simplest case, if m = 0,

$$\left(\frac{d \ln T}{d \ln p}\right)_{\text{rad}} = \frac{\tau}{4\left(\frac{3}{2}+\tau\right)}. \tag{2.82}$$

With an increase of optical depth this value tends to 0.25. It is apparent that the assumption that m = 0 does not correspond to reality. If it is assumed that m = 1/2, we obtain a dependence of the radiative gradient on optical depth shown in Table 20. In place of the dependence $\bar{\chi}=\bar{\chi}_0\tau^m$, it is possible to represent $\bar{\chi}$ through pressure, the empirical formula in this case will be:

$$\bar{\chi} = \bar{\chi}_0' p^n, \quad \nabla_{\text{rad}} = \frac{(n+1)\tau}{4\left(\tau+\frac{2}{3}\right)}. \tag{2.83}$$

In the surface layers it can be assumed that n = 1.8; the corresponding values for m = 0, m = 1/2 are also given in Table 20.

The derived values are characterized by the fact that they are either large, or they are equal to the adiabatic gradient at its normal value, obtained from formula (2.30), that is, 0.40. Thus, Table 20 can assist in obtaining at least a qualitative idea concerning the optical depth at which convection develops. The relationship between the radiative temperature gradient and the absorption coefficient is entirely obvious physically because where this coefficient is large, there are

Table 20

	Value $\bar{\tau}$						
	0	0.5	1	3	5	10	∞
m = 0	0.00	0.11	0.15	0.20	0.22	0.23	0.25
m = 1/2	0.00	0.21	0.30	0.40	0.44	0.47	0.50
n = 1.8	0.00	0.30	0.42	0.57	0.62	0.66	0.70

favorable conditions for an increase of the temperature gradient. The absorption coefficient attains a large value in the zone of ionization of hydrogen. When $\theta = \frac{5,040}{T} = 0.5$, the electron pressure is $p_e = 10^4$ dynes/cm^2. The Rosseland coefficient already is greater than 100. Ionization, therefore, is also an important factor in an increase of the radiative gradient.

We will now consider another aspect of the inequality representing the Schwarzschild criterion, that is, changes of the adiabatic gradient. The physical sense of the phenomenon is that the heat capacity (at constant pressure) increases appreciably in the course of ionization so that there is a change of c_p, and with it a change of the adiabatic gradient. The change in heat capacity is associated with the fact that new degrees of freedom arise during the ionization process, and each of them should obtain the value 1/2 kT; consequently, in the zone of its ionization a monatomic gas behaves as a multiatomic gas. It goes without saying that the picture is simplest quantitatively for the case of a single element (such as for a sun of pure hydrogen). In such a case, as was established long ago by Unsöld, the adiabatic gradient can be expressed through the degree of ionization

$$x = \frac{n}{N_0 + n}, \tag{2.84}$$

where n is the electron concentration; and, N_0 is the concentration of neutral particles, and can be expressed through its potential by the formula

$$\left(\frac{d \ln T}{d \ln p}\right)_{ad} = \frac{2 + x(1-x)\left(\frac{5}{2} + \frac{\varkappa}{kT}\right)}{5 + x(1-x)\left(\frac{5}{2} + \frac{\varkappa}{kT}\right)^2}. \tag{2.85}$$

Expression (2.85) becomes considerably more complex when there are two elements. The mean degree of ionization is introduced first. If the fraction of ionized atoms of stellar matter and the degree of ionization are v_1 and x_1, the fraction of such elements, which in a particular region of pressure and temperature either are not ionized at all or are ionized completely, and their degree of ionization is v_2 and x_2, the mean degree of ionization is expressed as:

$$x = \nu_1 x_1 + \nu_2 x_2. \tag{2.86}$$

In this case the adiabatic gradient assumes the form:

$$\left(\frac{d\ln T}{d\ln p}\right)_{\text{ad}} = \frac{\left(\frac{1+x}{x_1(1-x_1)} + \frac{\nu_1}{x}\right) + \nu_1\left(\frac{5}{2} + \frac{x_1}{kT}\right)}{\frac{5}{2}\left(\frac{1+x}{x_1(1-x_1)} + \frac{\nu_1}{x}\right) + \nu_1\left(\frac{5}{2} + \frac{x_1}{kT}\right)^2}. \tag{2.87}$$

Computations based on these formulas show that the adiabatic gradient is extremely sensitive to a change in the degree of ionization. At the same time, if it is taken into account that in the case of helium there is primary ionization, as well as secondary ionization at higher temperatures, computations on the basis of formulas of the type (2.83) and (2.87) will become too complex. Unsöld recommends that a different approach be used, that is, that an entropy diagram be used. It is known that in a reversible circular process, the entropy remains constant, that is

$$\bar{d}s = \frac{dq}{T} = 0. \tag{2.88}$$

Such processes are called isoentropic. It is obvious that an isoentropic process is at the same time adiabatic, since dq = 0.

Considering an arbitrary process, it can be represented graphically as follows: temperature at a linear scale is plotted along the x-axis, and the entropy values along the y-axis. Lines of equal entropy values—isoentropic lines—are represented by horizontal straight lines; and as was said above, they also will be adiabats, which is why such a diagram can be used in the final analysis for the formulated problem. We note, in passing, that it is easy to convert from entropy to potential temperature, that is, the temperature of a gas adiabatically reduced to a certain standard pressure by use of the formula

$$\Theta = \left(T\frac{P}{p}\right)^{\frac{\gamma-1}{\gamma}}, \tag{2.89}$$

where P is standard pressure (in meteorology a pressure of 1,000 mb usually is used), p is the given pressure, and, γ has its usual value. The relationship between entropy and potential temperature is given by the expression

$$\rho - s_0 = c_p \ln \frac{\theta}{\theta_0}, \tag{2.90}$$

where s_0 and θ_0 are certain initial values of entropy and potential temperature. Therefore, instead of laying off entropy at a linear scale along the y-axis, in meteorology the potential temperature at a logarithmic scale sometimes is plotted along this axis. On this graph the isotherms are represented by straight lines parallel to the y-axis and the isobars constitute a family of curves. Pressure p is a parameter constant for each curve. The form of such a thermodynamic diagram is shown in Figure 22. Here the dimensionality of entropy is 10 ergs/g ·deg (entropy also can be represented as cal/g·deg and as ergs/g·deg and as j/g·deg); in all these cases entropy is computed per unit mass. Unsöld computes it per atom; dividing the derived value by the Boltzmann constant k, he obtains a dimensionless number which he calls specific entropy. In computing this value he uses the formula (we will cite it here in a simplified form)

$$s_m = kN_m \left\{ \frac{5}{2} \ln N_m + \ln \frac{(2\pi M_m kT)^{3/2}}{h^3} + \ln u_m \right\}, \tag{2.91}$$

where N_m is the quantity of particles of an ideal gas; m in a unit volume; M_m is the mass of a particle; k is the Boltzmann constant; h is the Planck constant; T is temperature; and, u_m is the statistical weight of the ground state of the particular gas.

It is known that entropy is in fact a logarithm of the statistical weight. For example, the entropy of a particle $\frac{s_m}{N_m}$ consists of the entropy of the ground state plus that which is contributed by other factors. Total entropy is defined as the sum of the entropies of the component gases, that is, in this particular case—hydrogen, ionized hydrogen, neutral helium and once- or twice-ionized helium. After all s_m have been computed, these values are summed, and only then is entropy computed per single particle $\frac{s}{kN}$, where N is the total number of particles in a unit volume. The following two formulas were derived by Unsöld for two extreme cases (neutral gas and total ionization):

$$\frac{s}{kN} = 2.3026\,[6.080 + 2.5 \lg T - \lg p], \qquad (2.92)$$

$$\frac{s}{kN} = 2.3026\,[7.508 + 5.375 \lg T - 2.150\, p]. \qquad (2.93)$$

We note that in these computations, Unsöld assumes the composition of the sun to be as follows: 85 percent hydrogen and 15 percent helium. After computing the specific entropy the author constructs a diagram. A peculiarity of this diagram is that it is not temperature that is used as the independent variable, as on meteorological diagrams, but gas pressure. The reason for the selection of gas pressure as the independent variable is that in a study of processes in the convective zone, the temperature of the ascending and descending elements can be different from the temperature of the surrounding medium. As before, the isoentropic lines coincide with the adiabats, and like them will be straight lines, parallel to the horizontal axes (pressure axes), as on a meteorological diagram. In this case, however, it is isobars, not isotherms, which are parallel to the vertical axis. The isotherms will have a form quite similar to that of isobars on a meteorological diagram, but it is obvious that the parameter will now be T, not P. Since the entropy diagram in astrophysics covers a great range of depths, the diagram will have logarithmic instead of ordinary isotherms. An example of such a diagram has been reproduced in Figure 23. This graph shows bends in the isotherms, indicating zones of ionization of hydrogen and more weakly expressed zones of ionization of helium (single and double ionization). In order to obtain values of the adiabatic gradient it is necessary to carry out numerical differentiation along the adiabat.

2. Macroscopic theory

A peculiarity of convection under astrophysical conditions, and especially in the subphotospheric convective zone, is an interaction between convection and radiation. Radiative heat exchange has a very high value. Problems involved in the interaction between these two mechanisms of heat transfer have been studied by the Unsöld school (Ref. 37). We will now present the corresponding reasoning and conclusions. It should be noted that in many respects the investigation of this problem still is in its initial stages. Essentially, no allowance has been made for viscosity or Coriolis force, or especially for the electromagnetic factor, to such a degree that it is possible to take them into account when developing a model of the distribution of the principal characteristics with depth.

We will now consider the increment of the difference of vertical temperature gradients in the turbulent element and in the surrounding medium in the vertical segment h, that is, the value:

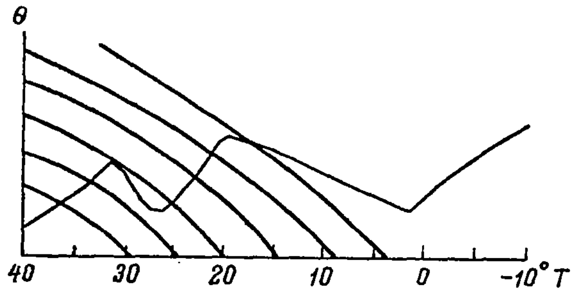


Figure 22. Simplified thermodynamic diagram. Potential temperature as a function of temperature; pressure is a parameter. The sloping curves are isobars.

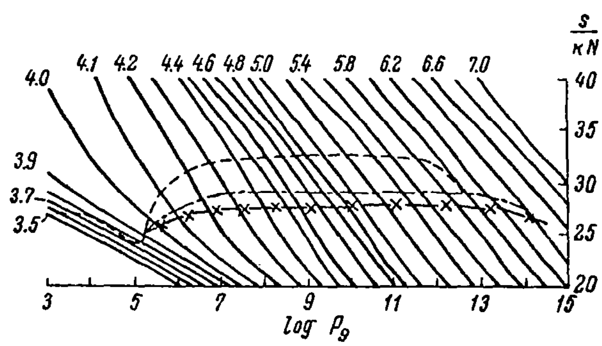


Figure 23. Entropy diagram in astrophysics (according to Unsold) for different relations I and H. Sloping curves are logarithmic isotherms.

$$\Delta\nabla T = h\left\{\left(\frac{dT}{dh}\right)' - \left(\frac{dT}{dh}\right)\right\}. \tag{2.94}$$

We will assume that $h = \frac{l}{2}$, that is, one-half the distance traveled by the turbulent element before its destruction; then

$$\Delta\nabla T = \frac{l}{2}\left\{\left(\frac{dT}{dh}\right)' - \left(\frac{dT}{dh}\right)\right\}. \tag{2.95}$$

We will introduce notations for the gradients (in Schwarzschild form): the radiative gradient is

$$\nabla_{\text{rad}}=\left(\frac{d\ln T}{d\ln p}\right)_{\text{rad}}$$

the gradient of true stratification is

$$\nabla=\left(\frac{d\ln T}{d\ln p}\right); \tag{2.96}$$

the gradient in a typical turbulent element is

$$\nabla'=\left(\frac{d\ln T}{d\ln p}\right)'; \tag{2.97}$$

the adiabatic gradient is

$$\nabla_{\text{ad}}=\left(\frac{d\ln T}{d\ln p}\right)_{\text{ad}}$$

We will now consider the difference in the gradients

$$\nabla'-\nabla=\left[\left(\frac{d\ln T}{d\ln p}\right)'-\left(\frac{d\ln T}{d\ln p}\right)\right]. \tag{2.98}$$

Since pressure in the element and in the surrounding medium is the same, then

$$\nabla'-\nabla=\frac{1}{d\ln p}\left[(d\ln p\,T)'-d\ln T)\right]. \tag{2.99}$$

After putting the barometric formula into logarithmic form, and differentiating for an isothermal atmosphere, we have

$$d\ln p=-\frac{\mu g}{\Re T}\,dh;$$

since $\frac{\mu g}{R}=\frac{1}{H}$, and where H is the height of the homogeneous atmosphere,

$$d\ln p=-\frac{dh}{H}\,. \tag{2.100}$$

After elementary transformations this makes it possible to rewrite (2.99) in the form

$$\frac{1}{H}(\nabla-\nabla')=\left[\frac{dT'}{T'dh}-\frac{dT}{Tdh}\right]. \tag{2.101}$$

Making the same assumption as when deriving the value of the adiabatic gradient, that is, assuming $T\approx T'$, we obtain (after multiplication of both sides by 1/2):

$$\frac{l}{2H}(\nabla-\nabla')T=\frac{l}{2}\left[\left(\frac{dT}{dh}\right)'-\left(\frac{dT}{dh}\right)\right]. \tag{2.102}$$

According to (2.95), the right-hand side of expression (2.102) is equal to $\Delta\nabla T$. Now taking the expression for the convective flux (2.57), and replacing $\Delta\nabla T$ by its value, we find that

$$\pi F_k=c_p\rho\bar{v}T\,\frac{l}{2H}(\nabla-\nabla'). \tag{2.103}$$

This formula could be called the astrophysical form of the expression for the convective flux, as the gradients included in it have been expressed in Schwarzschild form. It is also significant that it includes the height of the homogeneous isothermal atmosphere, and it is also important to establish the manner in which the convection velocity is expressed through this height. It follows from (2.56) that

$$\bar{u}^2=gl\frac{\Delta\nabla T}{T}. \tag{2.104}$$

Mean square velocity has been introduced here; averaging is accomplished for two velocity values—at the time of the "start" of the convective element, i.e., when h = 0, and, when h = 1. It is obvious that the factor 1/2 appears in this elementary averaging. Since $1/2H(\nabla - \nabla')$ $T = \Delta\nabla T$, it is found that

$$\frac{\Delta\nabla T}{T}=\frac{l}{2H}(\nabla-\nabla'). \tag{2.105}$$

After substituting (2.105) into (2.104), taking into account the factor 1/2, we obtain

$$\bar{u}^2=\frac{gl^2}{4H}(\nabla-\nabla'). \tag{2.106}$$

We now exclude $(\nabla - \nabla')$ from (2.103) and (2.106); then

$$\bar{u}^3=\frac{\pi F_k gl}{2c_p\rho T}.$$

Replacing $\rho T=\frac{\mu p}{\mathfrak{R}}$ (from the equation of state), we have

$$\bar{u}^3=\frac{\pi F_k gl\mathfrak{R}}{2c_p\mu p}. \tag{2.107}$$

It should be noted that this formula is correct when $\bar{u} < c_s$, where c_s is the speed of sound. This limitation is associated with the

fact that when $\bar{u} > c_s$, the conditions which served as a basis for derivation of formula (2.107) are not satisfied. The speed of sound is

$$c_s = \sqrt{\gamma \frac{\Re T}{\mu}}. \tag{2.108}$$

We then must establish the relationship between the different kinds of gradients. We note first that the total flux consists of the radiative and convective fluxes, that is

$$\pi F_k = \pi F_{rad} + \pi F_k. \tag{2.109}$$

When there is radiative equilibrium, the entire flux is transported by radiation; the convective flux is equal to zero, and the real stratification gradient is transformed into a radiative gradient. This can be written as follows (the detailed derivation is omitted):

$$\pi F = \pi F_{rad} = \frac{16}{3} \frac{\sigma T^4}{\varkappa \rho H} \nabla_{rad}. \tag{2.110}$$

Under real conditions, that is, when the energy is transported by both radiation and convection, the radiative flux can be expressed through the real stratification gradient (this derivation is omitted here due to lack of space). Then

$$\pi F_{rad} = \frac{16}{3} \frac{\sigma T^4}{\varkappa \rho H} \nabla. \tag{2.111}$$

It is obvious that the total flux can be represented through the effective temperature. Then, if the total flux is known, and also the effective temperature, it is possible to find the radiative gradient from the relation

$$\pi F = \sigma T_e^4 = \frac{16}{3} \frac{\sigma T^4}{\varkappa \rho H} \nabla_{rad}. \tag{2.112}$$

Thus, from formula (2.110) it is possible to find ∇_{rad}; from (2.111) - ∇; and, from (2.30) - ∇_{ad}.

We still must derive the expression for ∇'. Considering the turbulent element, the following expression is derived

$$\Gamma = \frac{\text{excess energy content}}{\text{radiation during lifetime}}. \tag{2.113}$$

Per unit volume the difference in the energy content of a turbulent element and a unit volume of the surrounding medium will be $c_p\rho\Delta\nabla T$ (this also is the excess energy content, or to be more precise, the density of this excess content). If in order of magnitude it is assumed that the linear dimensions of the turbulent element are l, and it is assumed that it has the form of a cube, its volume will be l^3, and its energy content will be

$$c_p\rho\Delta\nabla T l^3. \tag{2.114}$$

Two cases are possible:

1. The turbulent element is optically thin. This means that the coefficient of volume absorption, $\bar{\varkappa}\rho < 1$, is small. In this case l also is quite small, that is, the entire optical thickness $\bar{\chi}\rho l$ has a small value. The mixing length here is identified with the characteristic dimension of the element; in general, it is sufficient that they coincide in order of magnitude. Radiative heat exchange during the time $l/\bar{v}$ can be expressed as follows: the radiation coefficient is introduced from the relation $\bar{\epsilon}=\bar{\alpha}B$. Since the medium also radiates, there is radiative heat exchange between an element of the volume and the medium; therefore, ΔB is introduced in place of B, where Δ has the same sense as for T, that is, it reflects the change of B of the element in the process of its ascent. In an elementary solid angle in a unit time such an element of volume radiates $\bar{\chi}\rho\Delta B$.

During the time $l/\bar{v}$, the entire volume in the entire solid angle 4π radiates

$$4\pi\Delta B\bar{\varkappa}\rho l^3\frac{l}{\bar{v}}. \tag{2.115}$$

Taking into account that $B=\frac{\sigma}{\pi}T^4$, and substituting (2.114) and (2.115) into formula (2.113), we obtain (taking the finite difference, and shortening)

$$\Gamma_a=\frac{c_p\rho T\bar{v}}{16\sigma T^4}\frac{1}{\bar{\varkappa}\rho l}. \tag{2.116}$$

2. The turbulent element is optically thick. Here $\bar{\chi}\rho l \gg 1$. It is this case which is of particular interest for the subphotospheric layers of the sun. It is physically different from the first in that in the first case the entire volume radiated, since in the case of small values of the absorption coefficient there were no obstacles to radiation by the inner layers of the element, in the second case considered we need take into account only radiation from the surface. In this case

it is therefore necessary to determine the flux. The radiative flux can be expressed as follows (Ref. 37):

$$\pi F_{rad} = \frac{4\pi}{3} \frac{\Delta T^4}{\varkappa\rho l}. \tag{2.117}$$

In order to convert to the quantity of energy during the characteristic time $1/\bar{u}$ and from the surface $6l^2$ (since l is the edge of the cube), we multiply (2.117) by their product, and after substituting into the expression for Γ, we obtain

$$\Gamma_b = \frac{c_p \rho T \bar{u}}{32\sigma T^4} \varkappa\rho l. \tag{2.118}$$

In order for there to be convection, it is necessary to have the following relation between the gradients introduced into consideration:

$$\nabla_{rad} > \nabla > \nabla' > \nabla_{ad}. \tag{2.119}$$

The outermost terms of this inequality constitute the Schwarzschild criterion. The condition $\nabla > \nabla'$ should be satisfied in that only in this case will the temperature in the element decrease with height more slowly than in the surrounding medium. In order for there to be a convective flux through the surface of the element, it is necessary that $\nabla' > \nabla_{ad}$; finally since heat exchange with the external medium under real conditions always exists to a certain degree, the radiative gradient should be greater than the real gradient, that is, $\nabla_{rad} > \nabla$. Combination of these relations also gives the inequality (2.119).

In expression (2.118) the numerator obviously is proportional to the excess of the temperature gradient in the element above the gradient of the surrounding medium; the denominator should be proportional to the deviation from the adiabatic state, that is, $\nabla_{ad} - \nabla'$. In actuality, if the conditions were purely adiabatic, the denominator would contain only ∇_{ad}. Since there are expressions in both the numerator and denominator of (2.118), which have the dimensionality of energy, the proportionality factor should be the same for both the numerator and denominator. As a result, we therefore obtain the following dimensionless parameter for Γ:

$$\Gamma = \frac{\nabla' - \nabla}{\nabla_{ad} - \nabla}. \tag{2.120}$$

We will now write the expression for ∇_{rad} and ∇, and use the inequality $\nabla_{rad} > \nabla$; then

$$\frac{16}{3}\frac{\sigma T^4}{\varkappa\rho H}\nabla_{rad} > \frac{16}{3}\frac{\sigma T^4}{\varkappa\rho H}\nabla. \tag{2.121}$$

In order for this inequality to become an equality, we add to the right-hand side a term for the convective flux. Then, dividing both sides of expression (2.121) by $\frac{16\sigma T^4}{3\varkappa\rho H}$, and after substituting Γ_b from (2.118)—the optically thick element—we have

$$\nabla = \frac{\nabla_{rad} + 3\Gamma_b\nabla'}{1 + 3\Gamma_b}. \tag{2.122}$$

On the other hand, it is possible to use relation (2.120) for Γ. After excluding ∇ from (2.122) and from (2.120), we find an expression for ∇':

$$\nabla' = \frac{\nabla_{rad} + \Gamma_b(1 + 3\Gamma_b)\nabla_{ad}}{1 + \Gamma_b(1 + 3\Gamma_b)}. \tag{2.123}$$

Formulas (2.122), (2.123) and (2.118) for computation of the radiative gradient and the entropy diagram for finding the adiabatic gradient, and formula (2.106) for computation of square velocity, make possible the computation of subphotospheric stratification, taking into account convective transport, but under conditions of its joint effect with radiative transport. The most complete model of this type was developed in 1953 by Vitense (Ref. 45). This computation must be carried out by means of successive approximations, beginning with sufficiently deep layers, but those still accessible to observation.

Figures 24 and 25 show the results of computations by Vitense (Ref. 45). Figure 24 is the most important for a model of the subphotospheric layers. Gas pressure has been plotted along the x-axis and temperature along the y-axis. In addition, at the upper part of the figure we have indicated two scales of geometric depths, reckoned from the upper photosphere. One corresponds to the assumption that the mixing length is equal to one homogeneous atmosphere, and the other to two homogeneous atmospheres. However, this figure reflects temperature change with depth to relatively shallow depths (2,000-4,000 km), whereas according to Vitense the convective zone extends to depths of 66,000 km in the case in which $l = H$, and to a depth of 160,000 km when $l = 2H$. Figure 25 gives the change T^4, that is, values proportional to the energy of radiation in layers close to the photosphere. The solid and dashed

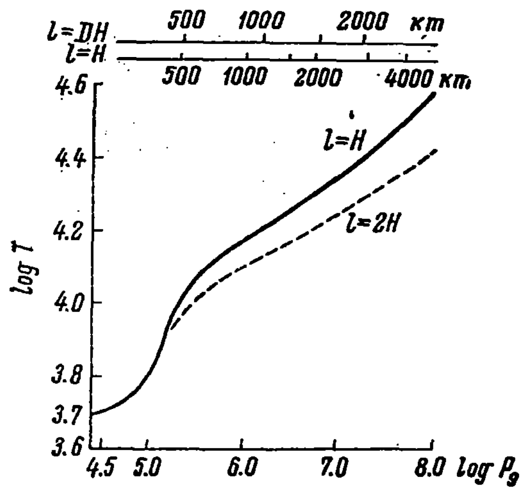


Figure 24. Change of temperature with depth in the convective zone, expressed through the logarithms of gas pressure, for two assumptions concerning the relationship between the mixing length and the height of the homogeneous atmosphere.

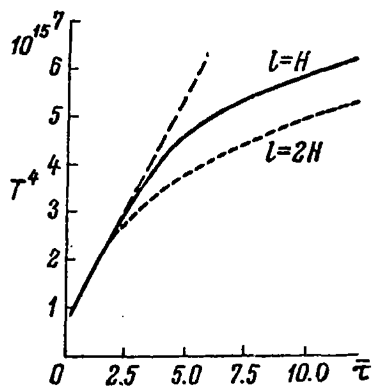


Figure 25. Change of value proportional to radiation energy with a change of depth in the convective zone.

curves represent the change of T^4 for l = H and l = 2H, respectively, and the straight dashed curve corresponds to conditions of radiative equilibrium.

Figure 26 shows the variation of the ratio of the convective flux to the entire flux with a change of depth (with gas pressure as an argument). We see that when log p_g = 5.5, which approximately corresponds to a depth of 500 km under the upper photosphere, virtually the entire transport of energy remains convective. Figure 27 shows the curve of the velocity of convective elements with depth. It can be seen that the velocity of these elements very rapidly attains a maximum value (about 3 km/sec when l = 2H, and about 2 km/sec when l = H) and

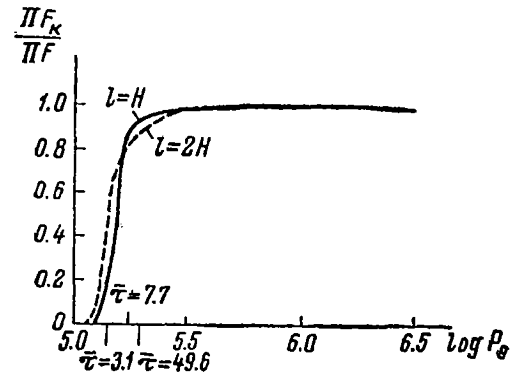


Figure 26. Change in the ratio of the convective flux to the total flux with change in depth in the convective zone

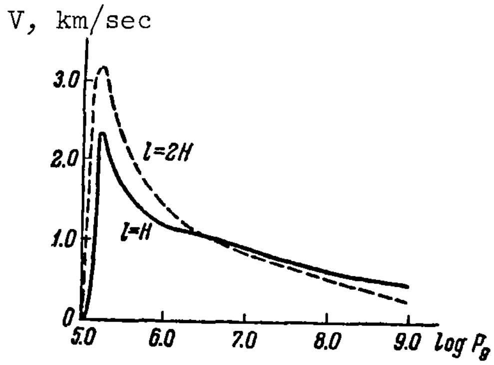


Figure 27. Change of vertical convective velocity with change of depth in the convective zone for two assumptions concerning the relationship between the mixing length and the height of the homogeneous atmosphere

then slowly drops off; but even at a depth of about 5,000 km under the photosphere it still constitutes about 0.5 km/sec. Figure 28 shows the change with depth of the difference in the temperature gradients in the turbulent element and in the surrounding medium. It can be seen that this change closely resembles the change of velocity of turbulent elements with depth.

Historical Development of Concepts Concerning the Convective Zone and Certain Astrophysical Conclusions

It was Unsöld in Ref. 46 who first indicated the possibility of convection in the outer layers of the sun. It was Unsöld who derived a considerable part of the principal relations of the theory of convection as applicable to astrophysical conditions. His results indicated that the convective zone occupies a space of very limited thickness, and is situated between the optical depths $\bar{\tau} = 3$ and $\bar{\tau} = 45$, and in the gas pressure range from $2.34 \cdot 10^2$ to $5 \cdot 10^3$.

According to later determinations, the gas pressure at those optical depths which the convective zone occupies according to Unsöld, is considerably greater; the values given by Unsöld apply to layers higher than the photosphere. If we proceed on the basis of modern concepts concerning the geometric depth of those optical layers which Unsöld considered, the thickness of the convective zone is considerably greater than that postulated by Unsöld. It is true that somewhat after the publication of his first study on this theme, he concluded that favorable conditions for the development of convection should be present between the optical depths $\bar{\tau} = 1.6$ to $\bar{\tau} = 170$, which according to his data correspond to a geometric thickness of about 500 km.

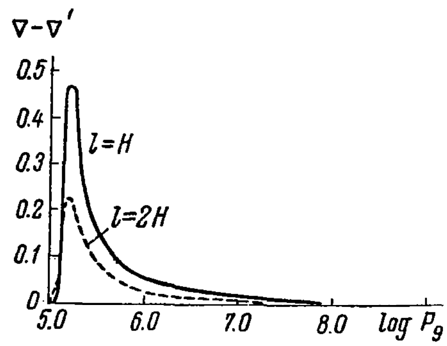


Figure 28. Change with depth of the difference of temperature gradients in the turbulent element and in the surrounding medium in the convective zone.

Siedentopf (Ref. 47) shared Unsold's opinion of that time, and assumed that the convective zone in actuality is relatively thin. It was this point of view on which Siedentopf based his original development of the convective theory of granulation. The role of convection in the principal manifestations of solar activity, such as sunspots, was regarded as unimportant. In any case, at that time it was assumed, and Unsöld shared this opinion, that the convective zone has no relation to solar activity due to its insignificant extent in depth in comparison with the solar radius. Waldmeier also assumed that the convective zone has no significant vertical extent (Ref. 48).

Biermann used a different approach to this problem (Ref. 49). He calculated that energy transport in the convective zone is accomplished for the most part by convection, not radiation. The Biermann hypothesis, therefore, should be regarded as the next approximation in comparison with the Unsöld and Siedentopf models. Biermann obtained a model of the distribution of the principal characteristics in the subphotospheric layers.

In the Biermann model the same gas pressure values correspond to lower temperatures; in addition, on the curve of the logarithm of temperature, there are no bends as there are in the Vitense model, corresponding to a zone of ionization. This can be attributed to the fact that in the Biermann model the zone of ionization is more extensive. The more rapid increase of temperature with pressure in the Vitense model may be attributed to the fact that the expenditure of energy on ionization in this model is less than in the Biermann model, and the increase of thermal energy with depth leads to a temperature increase greater than in the Biermann model.

Strömgren (Ref. 50) was a supporter of the concept of a deep convective zone. He proceeded on the basis of concepts very close to Biermann's point of view. Strömgren assumes that even when there is a temperature gradient close to radiative, the depth of the convective zone should be of the order of 20,000 km. However, if the temperature gradient is close to adiabatic, the depth of the zone should be considerably greater, a fact which Biermann pointed out before Strömgren did. It was Strömgren, however, who established the lower boundary of the convective zone for this case; it lies at a depth of about 200,000 km, where the temperature is about $2 \cdot 10^6$ degrees. This is the lowest boundary of the convective zone of all the boundaries determined by different investigators.

The problem of the convective zones obviously is not only of interest for solar physics; it is of broad astrophysical significance. The ratio of the convective flux to the total flux, expressed through effective temperature, can be represented as follows:

$$\frac{\pi F_k}{\pi F} = \frac{c_p T \rho \bar{u} \frac{l}{2H} (\nabla - \nabla')}{\sigma T_e^4}. \quad (2.124)$$

Convective transport loses its significance when $\pi F_k \ll \pi F$, that is, when

$$c_p T \rho \bar{u} \frac{l}{2H} (\nabla - \nabla') \ll \sigma T_e^4. \quad (2.125)$$

An analysis of this formula, made by Unsöld, revealed that under all circumstances $\nabla - \nabla' < 1$. Furthermore, $\bar{u}$ cannot exceed the speed of sound, since otherwise there would appear a serious obstacle for the rising turbulent element in the form of a sonic barrier. After substituting this value of critical velocity into (2.125), and taking into account that it follows from $c_s^2 = \frac{\gamma RT}{\mu}$ that $c_s = \sqrt{\gamma Hg}$, assuming $\gamma \approx 1$ and $l = H$, and also expressing ρ through the equation of state, after simple transformations we obtain

$$p < \frac{4}{5} \sigma T_e^4 \sqrt{\frac{\mu}{RT}}. \quad (2.126)$$

In Ref. 37 Unsöld computed data (Table 21) giving the lower limit of pressure at which there still is convective transport of energy. Effective temperature is the argument.

It is known that for stars of the main sequence, when $T_e < 8{,}000°$, $\log P = 5$, and at higher temperatures $\log p = 3$-4. Convection, therefore, does not influence the temperature stratification of stars of early classes. At a temperature below 6,000°, convective transport becomes increasingly important. It was established later that if there are traces of a convective zone in stars of classes A and F, the convective zone is very thin and is situated in the atmosphere of the star. On the other hand, in low-temperature stars the convective zone has a considerable extent.

Table 21

T	50000	20000	10000	5000	3000
p, dynes·cm^{-2}	10^8	$5\cdot10^6$	$4\cdot10^5$	$5\cdot10^4$	$9\cdot10^3$

We still must consider the following problems in this section: the supplementation of the "classical" theory of convective zones, representing the contributions of Böhm, Skumanich and others; problems relating to the influence of viscosity, rotation and electromagnetic factors on stellar convection; and, finally, the problem of models of the subphotospheric layers, "cementing" (a term sometimes used in hydrodynamics) of the convective zones to the inner zone of radiative equilibrium, since the way in which this "cementing" occurs to a considerable degree determines the structure of the convective zone itself.

We have already pointed out that the theory of the convective zone developed by Unsöld and his school cannot be considered perfect. After the appearance of the Vitense study in 1953 it could be assumed that an adequate approximation had been obtained. However, further investigations, especially the use of the results of a study of convection in general, yielded several new results.

In 1958 Böhm (Ref. 51) expressed the opinion that the usual assumption of equality of the height of the homogeneous atmosphere and the mixing length leads to difficulties; he therefore formulated the problem somewhat differently. According to Böhm there is a very unstable layer with a thickness of about 500 km lying over a far more stable layer with a thickness of about 60,000 km. The first can be called the upper, and the second the lower part of the convective zone. The difference between the actual and the adiabatic gradients $\nabla - \nabla_{ad}$ in the upper part of the zone, according to Böhm, is 100 times greater than in the lower part of the zone. Under these conditions the mixing length is equal to the thickness of the upper, most unstable part of the zone, that is, is about 500 km. Since the convective flux, all other conditions being equal, is proportional to the difference between the gradients, the energy transport by convection in the lower part of the convective zone should be far less than in its upper part. However, here it is necessary to take into account also the increase of T and ρ with depth.

If the model we adopted as a zero approximation (Table 17) is used as a zero approximation, an increase of density by a factor of 100 in comparison with its value at a depth of 500 km below the photosphere is attained only at a depth z = 40,000 km, where convective velocity should be considerably less than at a depth of 500 km. Thus, the convective flux should nevertheless decrease. An exception is the zone of maximum ionization, where it can again increase due to c_p.

The problem of the applicability of equation (2.57) for the convective flux under conditions of a considerable drop in density was questioned by A. I. Lebedinskiy (Ref. 52). The further development of convective theory as applicable to astrophysical problems also was pursued in the direction of clarification of the role of such a drop in

density. Taking this circumstance into account, Böhm and Richter determined at what length the disturbance wave would possess the maximum instability in the upper part of the convective zone; it was found to be 320 km (Ref. 53). Later these same investigators considered a three-layer model of the atmosphere: at top a zone of stability (photosphere), then a thin very unstable zone, and finally a zone of great extent and relatively low instability. The following conclusions could be drawn from this model of the structure of the subphotospheric layers: 1) the degree of instability increases rapidly with an increase of the wave number in those cases when the convective zone emerges at the surface of a star, that is, when the upper layer ceases to exist independently; 2) convection in the absence of lateral exchange increases more rapidly with time.

The problem of the convective zones was investigated recently by Biermann and his associates (Ref. 54). They calculated a series of models, varying the chemical composition, that is, the percentage content of hydrogen, helium and metals, and also the ratio of the mixing length to the height of the homogeneous atmosphere. This study contains a number of complex diagrams in which the following are given as a function of the logarithm of gas pressure: the logarithm of temperature, the adiabatic gradient in the Schwarzschild form, the logarithm of the ratio of the convective to the total energy flux, the mixing length, and the geometric depth. One of these diagrams shows the corresponding parameters obtained by Vitense in the above-cited study. This diagram has been reproduced in Figure 29.

In addition to such examples of more careful consideration of the interaction between convection and radiation, it is of considerable interest, as we have already mentioned, to take into account viscosity and other parameters. In this study Vitense states that allowance for such factors as turbulent friction, rotation and magnetic fields should lead to a decrease of the mixing length. At the same time, when Vitense did his work, the problem of the influence of viscosity on convection, as applicable to astrophysical conditions, still had not been developed. This problem was considered later by Skumanich and others (Ref. 55); they took into account viscosity under conditions when there are considerable changes of density between the upper and lower boundaries of the zone.

If we study the investigations made along these lines by the Unsold school, bearing in mind the different parameters entering into the Rayleigh criterion, it must be admitted that only heat conductivity has been taken into account properly. The interaction between convection and radiation was studied thoroughly, and radiation was taken into account to an appreciable degree, specifically in relation to radiative heat conductivity. No allowance was made even for ordinary viscosity, which enters into the Rayleigh criterion, not to mention turbulent

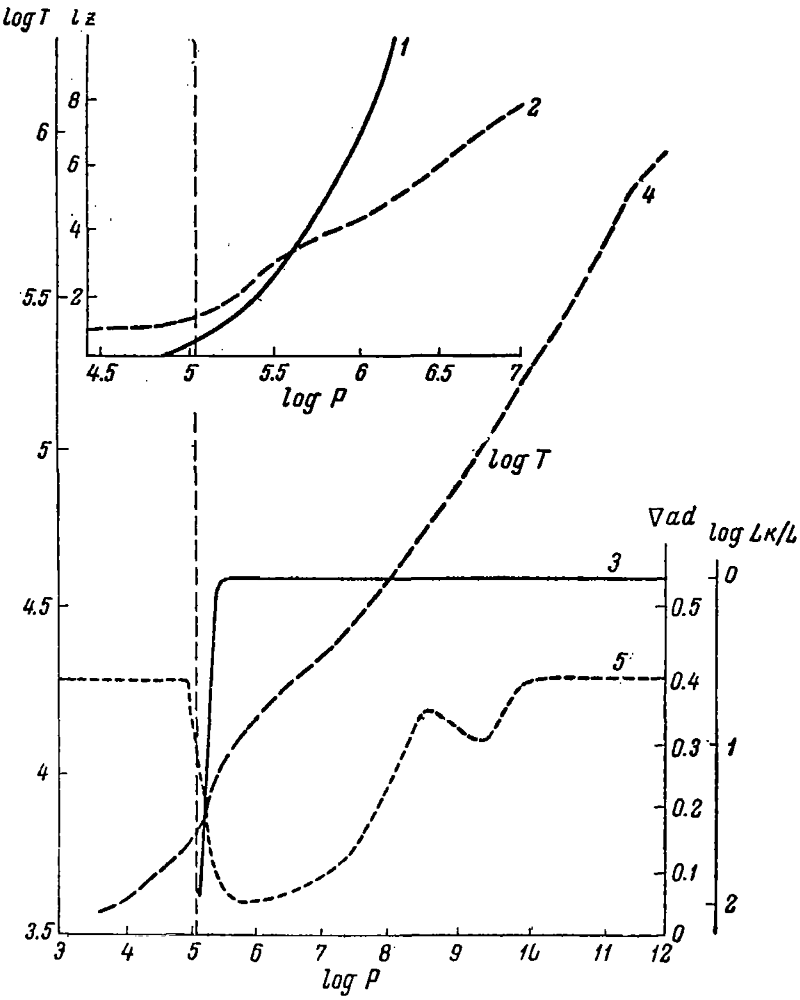


Figure 29. Change of various characteristics in the convective zone with change in depth; expressed through the logarithm of

gas pressure: 1, geometric depth, in 10^7 cm; 2, mixing length, in the same units; 3, logarithms of the ratio of the convective to total flux; 4, temperature; 5, adiabatic gradient.

viscosity. These problems require special consideration. Of course, it is most important to compute the convective flux because it is not possible to go along with the results obtained by Vitense on this point, as represented in Figure 26, where the convective flux relative to the total flux has been considered from the aspect of its change with depth,

gives a completely exhaustive result. With this in mind, however, it is important to consider not only the concepts which were developed by Böhm, but also the proper allowance for the factors mentioned above.

The role of viscosity also should be taken into account in the derivation of the Exner formulas, and when determining the velocity of convection. In general, it should be noted that the criteria developed by Rayleigh, Chandrassekhar, Taylor and others do not make it possible to determine directly and numerically the influence of the parameters in these criteria on specific values characterizing convection, such as its velocity, but are only a means for qualitative evaluation.

We will now proceed to the problem of the deeper subphotospheric layers and the connection ("cementing") with the model of the convective zone. We note first certain qualitative characteristics of the model of the structure of the convective zone according to Vitense. From these computations, it follows that the transition from ∇_{ad} to ∇_{rad}, that is, the equalization of the values of these gradients (which naturally means a cessation of convection, that is, it fixes the lower boundary of the zone) occurs in a jump at a gas pressure of $\log p_g = 11.44$, which, as already mentioned, with the most elementary assumption concerning the length of the mixing path (that is, when $l = H$), gives a depth beneath the photosphere of $z = 66{,}000$ km, or 1/10th of the solar radius.

On the other hand, the upper boundary of the zone is less clearly determined in this model and different authors indicate somewhat different optical depths for the beginning of convection—from $\bar{\tau} = 1.1$ to $\bar{\tau} = 3$. It is appropriate to note here that the presence of granules in the layers lying above the upper boundary of the convective zone, and that they are in radiative equilibrium (photosphere) at one time appeared to be a decisive objection against the convective theory of granulation.

The Unsöld school, and Vitense in particular, made particular efforts to overcome this elementary objection, and pointed out that the penetration of turbulent elements into the photospheric layers, with these elements experiencing dispersal during their rising in the zone of instability, occurs by inertia, that is, there is forced convection due to the initial impulse. However, if it is taken into account that viscosity, solar rotation and Chandrassekhar's Q parameter under real solar conditions always are different from zero, it will be impossible, a priori, to state that the turbulent elements can penetrate into the stable zone. Therefore, the same difficulties are encountered in the convective theory of granulation that were encountered earlier with

respect to this theory, before the investigations of the Unsold school. This, together with a number of other considerations, now forces us to give preference to the wave theory of granulation.

The convective zone is adjoined from below by an extremely thick region which is in a state of radiative equilibrium. Modern concepts concerning the internal structure of the sun exclude the existence of a convective core (Ref. 3), and, therefore, this inner region of radiative equilibrium reaches virtually to the most central layers. One of the most satisfactory models of the inner layers of the sun is a model proposed by Motz and Epstein (Ref. 56). It is true that in this model a small convective core has been retained which extends from the center to a distance equal to 8 percent of the solar radius. The presence of the small core is related to the fact that a certain, although very small role in the generation of energy is played by the carbon cycle. However, a valuable aspect of this model, as the authors point out in the foreword to their study, is that it also is applicable in a case when all the energy is formed due to the proton-proton reaction (obviously, they mean that such an assumption does not cause additional changes in the model for that part of the sun considered to be in radiative equilibrium).

The composition of solar matter is assumed to be as follows: hydrogen, 93.1 percent; helium 6.7 percent; metals and others, 0.2 percent. Under these assumptions the pressure, temperature and density in the layer falling between 200,000-603,000 km from the center of the sun are represented in Table 22. These results were correlated by Biermann and Temesvary (Ref. 57) with the model of the convective zone developed by Vitense. Their point of departure was a critical examination of the theories held by Schmeidler concerning stationary forced convection, which during the lifetime of the sun should be propagated from the convective zone to the most central regions. It was demonstrated that such a theory is untenable. If there was such a possibility, the temperature and density in the interior of the sun would be inadequate for the occurrence of thermonuclear reactions. In addition, they demonstrated that the energy necessary for propagation of the zone of instability into the deeper layers should be commensurable with the energy transported to the sun's surface from the inner layers.

It obvious that there is no source for such an influx of energy into the peripheral layers from the outside. Biermann and Temesvary then investigated the behavior of the logarithmic gradient of radiative pressure and demonstrated in this manner that near the lower boundary of the convective zone, the difference between such a gradient and the adiabatic gradient does not exceed 10^{-5}, and apparently is even considerably less.

Table 22

log r	log p	log T	log ρ
10.78	10.75	5.57	-3.02
10.74	11.92	5.81	-2.09
10.70	12.74	5.98	-1.44
10.66	13.37	6.11	-0.94
10.62	13.88	6.24	-0.56
10.58	14.29	6.35	-0.26
10.54	14.65	6.44	0.01
10.50	14.97	6.51	0.26
10.46	15.26	6.57	0.49
10.42	15.52	6.62	0.70
10.38	15.77	6.67	0.90
10.34	15.98	6.72	1.06
10.30	16.18	6.77	1.21

For our purposes it is of the greatest interest to note the correlation which these investigators made between the Vitense model of the convective zone and the Motz-Epstein model. This result is represented in Figure 30. Pressure, as usual, has been plotted along the x-axis; for all practical purposes, of course, this is gas pressure. The ratio of the total pressure (gas + radiative) to gas pressure for the sun (β) is approximately 0.99 (Ref. 5). Biermann and Temesvary consider three values $1 - \beta$: 10^{-2}, 10^{-3} and 10^{-4}. In the case of such constant values the dependence of log T on log p is represented by a straight line. In this case the molecular weight is assumed to be 0.7, whereas in the Motz-Epstein model it is 0.52.

The numbers on the graph in Figure 30 correspond to the relative radius $r/r_\odot$. This same figure shows the polytropic line with the index $n = 3/2$, having the same radius and the same mass as the sun; in this case the dependence of T on p follows from the Emden theory of polytropic spheres (Ref. 58). The polytropic curve has been considered in the range of gas pressure from log p = 11 to log p = 16, since in the layers closer to the surface the approximation of a polytropic sphere does not apply. The solid curve gives the Motz-Epstein model, beginning with $r/r_\odot = 0.87$.

In layers deeper than $r/r_\odot = 0.15$ or $\log p_g = 17$, the model cannot be applied, since they again lie beyond the limits of the radiative envelope in the region of the small convective core (see above). The

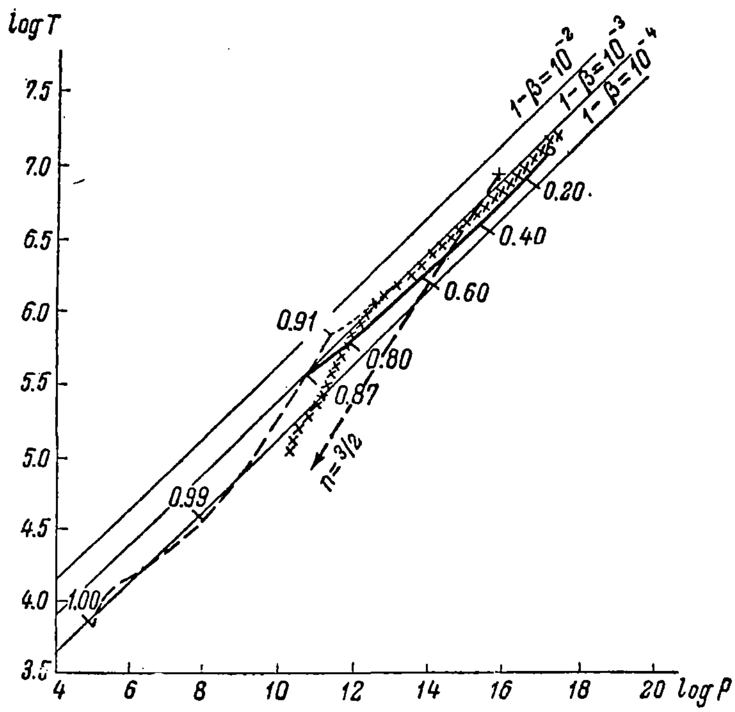


Figure 30. Biermann-Temesvary model; tie-in between the Vitense and Motz-Epstein models

lower boundary of the polytropic approximation does not even reach to this value. The Vitense curve can be seen with all its characteristic details. It begins with log p_g = 5, that is, z = 400 km, and extends to log p_g = 11.44—the lower boundary of the convective zone.

At the same time, it can be seen from Table 22 that the zone of radiative equilibrium according to Motz and Epstein begins at log p_g = 10.75. In this respect, the "cementing" of the convective and radiative models, accomplished by Biermann and Temesvary, cannot be regarded as entirely satisfactory. There can be two variants: first, it can be assumed that convective conditions are maintained to log p_g=10.74, and then it is necessary to use the Motz-Epstein model; second, it is possible to take the lower boundary as determined by Vitense, that is, log p_g = 11.44, and then use the Motz-Epstein model.

In Figure 30 Biermann and Temesvary interpolate a transitional zone, carrying the interpolation curve to a straight line corresponding to $\beta = 10^{-2}$. Vitense's point of view that the transition to conditions of radiative equilibrium is quite sharp is not refuted by the data in Figure 30. In this same figure the Weyman model (Ref. 59) has been given for comparative purposes. This model, like the Searce model (Ref. 60), and the model developed by Schwarzschild, Howard and Harm (Ref. 61), are not homogeneous. They take into account both the presence of a convective zone and an extended zone of radiative equilibrium.

The upper layers of the Weyman convective zone, however, are not considered, and the model begins with $\log p_g = 10.063$, which corresponds to a relative radius $r/r_\odot = 0.98$, or a depth beneath the photosphere, assuming a radius corresponding to an optical depth $\tau = 1$, equal to $6.96 \cdot 10^{10}$ cm, $z = 16{,}000$ km. However, the entire convective zone in the Weyman model extends to a depth of 140,000 km. The curve of log T as a function of $\log p_g$ is steeper in the convective zone than in the zone of radiative equilibrium. For this region of the Weyman curve, for which there are data on the Vitense convective zone, log T in the Vitense model lies parallel to its value in the Weyman model, but the same values $\log p_g$ in the model developed by the Unsöld school correspond to higher values than in the Weyman model. The Motz-Epstein curve begins where, according to the Weyman model, there is still part of the convective zone.

In the zone of radiative equilibrium, the Weyman and Motz-Epstein curves diverge from one another at a small angle, converging somewhat with an increase of $\log p_g$ (the Motz-Epstein curve is somewhat steeper than the Weyman curve). In general, in the zone of radiative equilibrium the temperature changes as a function of pressure more slowly than in the convective zone.

Finally Adopted Model

As the basis for the final model, we assumed the structure of the convective zone as determined by Vitense, since only in this model are there values directly from the subphotospheric layers. The zone of radiative equilibrium was taken in accordance with the Motz-Epstein model. "Cementing" was accomplished somewhat differently than by Biermann and Temesvary—immediately after the value log T corresponding

to log $p_g = 11.44$, we took the value for the zone of radiative equilibrium (Motz-Epstein); this makes the lower boundary of the convective zone sharper. In this sense it corresponds better to the initial idea of a convective zone as formulated by Vitense. In addition to the curve log T, the graph (Figure 31) also shows the curve ρ for atomic weight 0.7 (convective zone) and 0.52 (zone of radiative equilibrium).

Hereafter, in this chapter and in those which follow, when it is necessary to make preliminary computations, we have taken into account the data obtained from use of this model. The values for viscosity (gas kinetic and radiative), heat conductivity and gas kinetic electrical conductivity, cited in Tables 17 and 18 for the corresponding depths, strictly speaking, should be recomputed in accordance with the final model. However, taking into account that the computations made hereafter in this and in the chapters which follow have been made with an accuracy to an order of magnitude, we will pass over the corresponding recomputation.

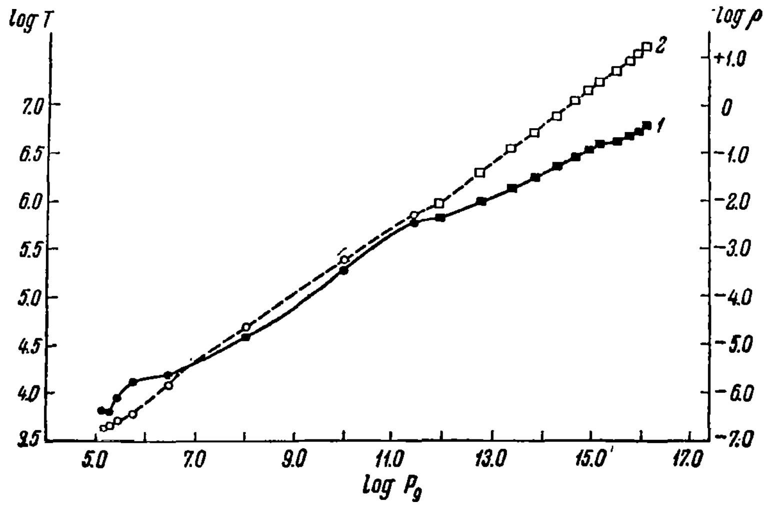


Figure 31. Final model of the subphotospheric layers: 1, most probable distribution of temperature in the subphotospheric layers; 2, density change with change in depth as a function of the logarithm of gas pressure

Section 3. Dynamics of Subphotospheric Layers

Distribution of the Angular Velocity of Rotation of the Sun With Depth and Rotational Instability

In this section we will discuss large-scale motions in the subphotospheric layers. Convection, velocity and other characteristics which were considered in the preceding sections apply to motions of a lesser scale. The same should be said also of such manifestations of solar activity as sunspots, etc. Such a differentation of the motions of higher rank and the scale of manifestations of solar activity was proposed by Rosseland in the important study, "The Theory of Rotating Stars" (Ref. 62). The known large-scale motions almost exclusively involve meridional currents, and the latter are associated closely with solar rotation. In connection with this exposition it is necessary to begin with certain additions to the theory of solar rotation.

As is well known, stars consisting of an ideal gas (and not of a degenerate gas, like white dwarfs) rotate in such a way that their angular velocity is a function both of the distance of a particular point from the axis of rotation and the distance from the equatorial plane (Ref. 9). It is true that with respect to the dependence of angular velocity on astrographic latitude there are other hypotheses which do not relate this dependence to the properties of an ideal gas under radiative conditions.

With respect to the dependence of angular velocity on the distance of a particular point to the axis of rotation, it certainly is difficult to propose a different explanation than the generally accepted mechanical and hydrodynamic mechanisms. Historically, the first was proposed by Jeans (Ref. 63) somewhat earlier than the second (that of Rosseland). We will first discuss the second, and then return to the first, since on its basis we have made certain additional computations, using the model of the subphotospheric stratification considered in the preceding section.

If it is assumed that a star consists of a homogeneous, incompressible fluid (this assumption is very far from real conditions), then viscosity need not be taken into account to satisfy the Helmholtz-Kelvin condition concerning the constancy of circulation, that is

$$\frac{dC}{dt}=0. \tag{2.127}$$

If we select as the circuit a circle, concentric with the axis of rotation and having the radius r, when the angular velocity on its periphery is ω, circulation obviously can be written as follows:

$$C = 2\pi r^2 \omega.$$

We will now assume that the particles making up the circuit are displaced by some value Δr, with the result that the radius of the circuit becomes r_1. As a result of the constancy of C there also should be a change of angular velocity, that is, the following relation should be satisfied

$$2\pi r^2 \omega = 2\pi r_1^2 \omega_1.$$

Hence

$$\frac{\omega}{\omega_1} = \frac{r_1^2}{r^2}. \qquad (2.128)$$

There are very many doubtful assumptions in this derivation. We have already discussed one of them, directly associated with the concept of satisfaction of the Helmholtz-Kelvin theorem, although stellar, and especially solar gas should be considered as a baroclinic fluid for which, as is well known, a different circulation theorem is correct, that is, the Bjerknes theorem. Under these conditions the right-hand side of (2.127) no longer is equal to zero. It is true that this consideration does not play an important role in the study of a case of rotation as circulation; it is important in essence when considering meridional currents, also associated with circulation, since in this case there is a drop of pressure and temperature at the boundaries of the circuit.[1] But the presence of meridional currents nevertheless makes inexact the assertion that the right-hand side of (2.127) is equal to zero, since they lead to a change of the circuit along which there is rotational motion with time (the resulting picture is similar to that which is observed in the earth's trade winds).

In general, the right-hand side of (2.127) should be equal to $-2\omega \frac{d\Sigma}{dt}$, where Σ is the area of the circuit. It is therefore natural that the true law of change of ω with depth will be more complex than the simple relation (2.127). The law (2.127) is therefore correct only to the degree to which the assumptions made are correct and to the extent to which it is possible to neglect the deviation of the right-hand side of (2.127) from zero. The hypothesis is facilitated to a certain degree in that meridional currents are slow in comparison with solar rotation; therefore, the right-hand side of (2.127) should be an insignificant value.

---

[1]Such meridional currents in general cannot exist in a baroclinic fluid.

Jeans, who considered the problem of the dependence of the angular velocity of solar rotation on depth before Rosseland did, approached the problem on a completely different basis. The basic idea was as follows: radiation (that is, transported radiant energy) is equivalent to some mass, and, therefore, radiation also can transport some moment of momentum. It is obvious that the latter has a small value in the layers of stars close to the center, and increases toward the periphery as a result of increase of distance from the center.

If we consider a certain volume within a star (sun), the radiation will convey to it the deficit of angular momentum from the inner layers, and viscosity (friction) will cause the onset of an excess of angular momentum from the periphery. It is found that under equilibrium conditions there will be a dependence of angular velocity of rotation on distance to the center, and after certain simplifications, this dependence assumes the same form as that given by formula (2.128). Figure 32 shows the mathematical expression of these relations. The surface AB, cut by an elementary cone with the solid angle $\sin\theta d\theta d\varphi$ (here φ is the azimuthal angle) will be $r^2 \sin\theta d\theta d\varphi$. If the radiation flux passing through AB is denoted H, the energy transported through AB in a unit time obviously will be $Hr^2 \sin\theta d\theta d\varphi$. The corresponding mass will be $\frac{H}{c^2} r^2 \sin\theta d\theta d\varphi$. After denoting by ω the angular velocity of the element AB in its rotation around the sun's axis (or the axis of a star), we obtain for the linear velocity of rotation $\omega r \sin\theta$, and this means that the moment of momentum transported by the radiation intersecting AB in a unit time will be

$$\frac{H}{c^2} r^4 \sin^3\theta d\theta d\varphi\omega. \tag{2.129}$$

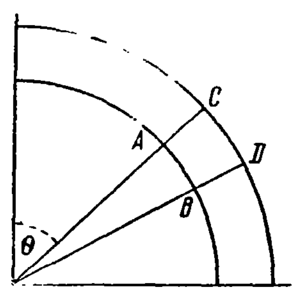


Figure 32. Diagram explaining the derivation of the Jeans relation

This is the quantity of angular momentum which every second will enter into the volume ABCD; after introducing r + dr in place of r, we will obviously obtain the quantity of angular momentum which emerges every second from the volume ABCD through the surface CD. The net loss of momentum per second, therefore, will be

$$\frac{d}{dr}\left(\frac{H}{c^2} r^4 \sin^3 \theta d\theta d\varphi \cdot \omega\right) dr. \tag{2.130}$$

Because of the viscosity in the considered volume, there will be an onset of an excess of angular momentum from the outer layers. A momentum of $\eta r \sin\theta \frac{d\omega}{dt}$ (here η is the viscosity coefficient) will pass inward through the unit surface AB, and the total angular momentum will be

$$\eta r^4 \sin^3 \theta d\theta d\varphi \frac{d\omega}{dr}. \tag{2.131}$$

In exactly the same way, it is possible to estimate the influx of momentum through CD. The net increase of angular momentum, therefore, will be

$$\frac{d}{dr}\left(\eta r^4 \sin^3 \theta d\theta d\varphi \frac{d\omega}{dr}\right) dr. \tag{2.132}$$

The change of momentum in each elementary volume can be considered the algebraic sum of (2.130) and (2.132), that is

$$\rho r^2 \frac{d}{dt}(r^2\omega) = \frac{d}{dt}\left[\eta r^4 \frac{d\omega}{dr} - \frac{H}{c^2} r^4 \omega\right]; \tag{2.133}$$

since $\rho r^2 dr \sin\theta d\theta d\varphi$ is the mass of the elementary volume, after shortening we obtain the left-hand side of (2.133).

Carrying this reasoning further, Jeans arrived at a slightly different expression:

$$\rho r^2 \frac{d}{dt}(r^2\omega) = \frac{\partial}{\partial r}\left[\eta r^4 \frac{d\omega}{dr} - \frac{3}{5}\frac{H}{c^2} r^4 \omega\right]. \tag{2.134}$$

In the inner regions of a star both terms of the right-hand side are small and, therefore, it can be assumed that there $\frac{d}{dt}(r^2\omega) = 0$. In the outer layers, in the case of small ρ the value ω changes rapidly, and this change leads to a stationary state in which the left-hand side of equation (2.134) becomes equal to zero. In this case, that is, when

$$\frac{\partial}{\partial r}\left[\eta r^4 \frac{d\omega}{dr} - \frac{3}{5}\frac{H}{c^2} r^4 \omega\right] = 0$$

we should have

$$\eta r^4 \frac{d\omega}{dr} - \frac{3}{5}\frac{H}{c^2} r^4 \omega = f(\theta), \tag{2.135}$$

where f is an arbitrary function of θ, which is colatitude. If we introduce here an expression for radiative viscosity (since, as follows from Tables 17 and 18, radiative viscosity in a rather thick subphotospheric layer exceeds the gas kinetic viscosity) and express this radiative viscosity through the radiation flux, according to Jeans we will obtain the following:

$$\eta = -\frac{HT}{10c^2} \Big/ \frac{dT}{dr}. \tag{2.136}$$

Substituting (2.136) into (2.135), we obtain

$$\frac{r^4 H}{c^2}\left(-\frac{1}{10}\frac{T}{\frac{dT}{dr}}\frac{d\omega}{dr} - \frac{3}{5}\omega\right) = f(\theta). \tag{2.137}$$

Hence

$$\omega - \frac{A}{r^2} = -\frac{1}{6}\frac{T}{\frac{dT}{dr}}\frac{d\omega}{dr}, \tag{2.138}$$

where A is a function of θ. In the first approximation we find

$$-\frac{1}{6}\frac{T}{\frac{dT}{dr}}\frac{d\omega}{dr} = \frac{T\omega}{3r\frac{dT}{dr}}. \tag{2.139}$$

Using (2.139) as a first approximation for ω, Jeans finds that

$$\omega\left(1 - \frac{T}{3r\frac{dT}{dr}}\right) = \frac{A}{r^2}. \tag{2.140}$$

Evaluating the values entering into this expression, he finally arrives at the following

$$\omega = \frac{A}{r^2}. \tag{2.141}$$

Thus, in the first approximation we actually obtain the same law as obtained by Rosseland from approximate hydrodynamic considerations. This coincidence gave Rosseland a basis for postulating a similarity of

neutral equilibrium for the inner layers of a star: if the laws derived on the basis of hydrodynamics and mechanics, with radiation taken into account, contradict one another this would be a factor inciting instability. At the same time, there is an appreciable difference in the very premises of the Rosseland hydrodynamic approach and the Jeans "dynamic" approach. Whereas in the hydrodynamic approach viscosity is not taken into account, and cannot be taken into account, because if viscosity is present the Kelvin-Helmholtz theorem would not be satisfied (viscous forces would enter into the right-hand side of the expression for the derivative of circulation), the Jeans study postulates viscosity as the only possible mechanism (without touching on what kind of viscosity is involved) for the transport of excess angular momentum from the outer to the inner layers. Jeans very successfully likens this effect to the effect of a flywheel. At the same time he points out that the outer layers, as a result of low density, do not have a very high angular momentum, despite the considerable velocity of their rotation.

The distribution of the angular velocity of solar rotation with depth is of considerable importance in the problem of the stability of the subphotospheric layers. It apparently was Rosseland in the study cited above who was the first to formulate the problem of rotational instability as applied to astrophysical problems, of course, since in hydrodynamics, in general, the corresponding theory was developed by Taylor (Ref. 41).

If with the movement of the contour we obtain an excess of angular velocity (for a particular distance from the axis of rotation), the motion is unstable due to an excess of centrifugal force. It is found that with an inadequacy of angular velocity (once again, for a particular distance from the axis of rotation) the motion will be stable and the developing perturbations will lead to a similarity of wave motions. As also noted by Rosseland, all of this is correct under conditions of neutral thermal equilibrium. However, if there is thermal stability it first is necessary for rotational instability to overcome its reserve. In theory it also is possible to have stabilization of a layer of rotation which before was thermally unstable. In the study of this problem it is common to use as a point of departure the differential form of relation (2.128) and use the following criterion (Ref. 10):

$$\frac{d(\omega r^2)}{dr} \geqslant 0. \qquad (2.142)$$

This is the case of stable equilibrium. On the basis of experiments with rotating cylinders, Taylor has shown that by appropriately selecting angular velocities it is possible to obtain a full similarity of convection without changing thermal stratification, and this can be done by using the phenomenon of rotational instability as a point of departure. In Ref. 64 Wasiutynski developed a detailed theory applicable to astrophysical problems.

The presence of an extended convective zone with its characteristic structure on the sun naturally exerts an appreciable influence on the distribution of angular velocity with depth. The most important fact is that it is turbulent viscosity which is of decisive importance in this zone, whereas in the below-lying zone of radiative equilibrium, it may not be. The problem of the method of transport of energy and angular momentum also is of great importance.

We will, therefore, formulate the following problem: an attempt will be made to compute the distribution of the angular velocity of solar rotation with depth in the convective zone, taking into account turbulent viscosity, and also the combined radiative and convective transport of angular momentum. It should be noted that the transport of angular momentum by convection, all other conditions being equal, is a more effective mechanism than its transport by radiation. In actuality, in the case of transport of the moment of momentum by convection, in order to obtain the mass from the energy, the energy must be divided by $v^2/2$, whereas in the case of transport by radiation it is necessary to divide by c^2.

We have seen that even in that place in the convective zone where according to Vitense the velocity of the ascending convective currents is maximum, it does not exceed several kilometers per second, which of course is insignificantly small in comparison with the speed of light. In order to introduce turbulent viscosity into consideration, we will begin with equation (2.135), but instead of radiative viscosity, which is introduced by Jeans, who obtained equation (2.137), we will introduce turbulent viscosity by using the relation

$$A = K_{*}\rho. \tag{2.143}$$

Here A is the exchange coefficient. We immediately will attempt to estimate what such an introduction yields in the sense of change of the orders of magnitude of the values entering into (2.135). For example, for a depth of 10^8 under the photosphere, the radiative viscosity has the order of 10^{-5}. In order to find the exchange coefficient we multiply the kinematic turbulent viscosity by the density, which at this depth is of the order of 10^{-6}-10^{-5}. The coefficient of kinematic turbulent viscosity for motion of a scale commensurable with rotation will be at least 10^{12}. Turbulent viscosity, therefore, will be many orders of magnitude greater than radiative viscosity. The transport of positive

angular momentum from the outer into the inner layers, therefore, will be considerably more effective than in the case of radiative viscosity.

Under the above conditions, for a state which is stationary with respect to angular momentum, that is, for satisfaction of its balance, the angular velocity of rotation should not increase with depth, but should instead decrease. It must be remembered, however, that due to convection the transport of negative angular momentum from the inner to the outer layers also will be more effective than in the case of its transport by radiation. Therefore, the final solution of the problem of how the angular velocity of solar rotation in the convective zone is dependent on distance to the axis can be obtained only after solution of the following equation:

$$K_s \rho(r) r^4 \frac{d\omega(r)}{dr} - \frac{2\pi F_k}{v^2(r)} r^4 (\omega) r = f(\theta), \tag{2.144}$$

where ω is the angular velocity of solar rotation; r is the radius vector of the considered point; K_s is the coefficient of kinematic turbulent viscosity; πF_k is the convective flux; ρ is density; v is the vertical velocity of the convective elements; and, $f(\theta)$ is an arbitrary function of the polar distance θ.

Replacing $f(\theta)$ by such a function as the Faye function $a-b\cos^2\theta$ and considering conditions at the sun's equator, we obviously obtain instead of $a-b\cos^2\theta$ an expression for a, that is, the equatorial velocity, which can be denoted ω_0. The equation (2.144) can be integrated numerically, but we will use a simpler approach. We will compare the equation (2.144) with the similar Jeans equation (2.135). The form of these equations is completely identical, but radiative viscosity enters into the Jeans equation and the transport of momentum is accomplished by radiation, whereas in equation (2.144) the momentum is transported by convection and the viscosity is turbulent.

Therefore, for a qualitative solution of the problem of how the angular velocity of solar rotation will change with depth under the conditions prevailing in the convective zone, it is sufficient to consider in what relation is the change of both the loss of the deficit of angular momentum from deep layers and the transport of its excess from the relatively outer layers. The energy flux in the photosphere, as is well known, is $6\cdot 10^{10}$ ergs$\cdot$cm$^{-2}\cdot$sec^{-1}. If at a depth of 10^8 cm below the photosphere, which we considered above, the energy also is transported by radiation, the flux would only be slightly greater than in the photosphere; the increase would be attained due to the curvature of the

layers, whereas under conditions of radiative equilibrium and plane-parallel layers the flux remains constant. A value of 10^{11} ergs$\cdot$cm$^{-2}\cdot$sec^{-1} even will be somewhat too high, taking into account that 10^{8} cm is only a very small fraction of the solar radius. The ratio of the flux to the square of the speed of light will be 10^{-10} cm$^{-2}\cdot$sec^{-1} ("flux of mass"). Radiative viscosity at this depth is $\approx 5\cdot 10^{-5}$g$\cdot$sec^{-1}. If the same energy flux was caused by convection, the "flux of mass" in the case of a convective velocity at the depth of $\approx 10^{5}$ cm$\cdot$sec^{-1} will be ≈ 10 g$\cdot$cm$^{-2}\cdot$sec^{-1}.

Turbulent velocity can be estimated approximately, using as a point of departure that the characteristic scale of the convective zone (with respect to its vertical extent) is $\approx 6\cdot 10^{9}$ cm and the mean velocity of convection is close to 10^{4} cm$\cdot$sec^{-1}. Then $K_s \approx 5\cdot 10^{13}$ cm$^{2}\cdot$sec^{-1}, and taking into account that density at this depth is $5\cdot 10^{6}$ g$\cdot$cm^{-3}, we obtain $K_s\,\rho \approx 10^{8}$. Thus, for the considered depth the relation of the first terms of equations (2.144) and (2.135) is $\approx 10^{12}$, and of the second terms, $\approx 10^{11}$.

Thus, under the conditions prevailing in the convective zone and the turbulent viscosity in it, the increase of angular velocity with depth should occur more slowly than indicated by the law of inverse squares. Even if we consider the turbulence associated with horizontal motions of the scale of 10^{11} cm (rather than that associated with vertical motions) and the velocity of circulation is $\approx 10^{2}$ cm$\cdot$sec^{-1}, the coefficient of kinematic turbulent viscosity will be $\approx 10^{13}$ cm$^{2}\cdot$sec^{-1}. This gives almost the same as given above.

It can therefore be expected that in the convective zone an increase of angular velocity with depth occurs more slowly than indicated by the law of inverse squares. The result is evidence in support of rotational stability of the convective zone, although to use the terminology employed by Wasiutynski (Ref. 11) it is gravitationally unstable.

It is interesting to note that in one of his recent studies (Ref. 64) Wasiutynski lends support to the rotational instability of the subphotospheric layers and assumes that the only force opposing this is the electromagnetic forces appearing at great depths beneath the photosphere (approximately at the base of the convective zone), where there is a strong magnetic field.

Meridional Currents and Circulatory Instability

The problem of the nature of the differential rotation of the sun, that is, the dependence of the angular velocity of its rotation on heliographic latitude, has an extremely long history. Without entering into a detailed historical excursion, for lack of space we note only that at the present time various monographs give different interpretations of this phenomenon. For example, Cowling (Ref. 8), beginning with an exposition of isorotation, i.e., the property of retaining the same angular velocity along a particular magnetic line of force, then comes to the conclusion that this hypothesis is untenable and is inclined to interpret solar differential rotation as some cosmogonic relict, that is, as a trace of the epoch when the sun was in its formation.

This point of view, incidentally, is not new; it was expressed as early as the end of the last century by Wilsing and Vil'chinskiy (Ref. 65), and will scarcely withstand criticism. In addition to those arguments against the theory of isorotation which are cited by Cowling, it is possible to point out the obvious weakness of the general magnetic field of the sun and the circumstance that isorotation at best could be a means for maintaining an already established law of rotation; the origin of this property must be sought elsewhere.

Still another hypothesis on the problem of the nature of differential solar rotation relates this phenomenon to the Jeans law, on the one hand, and on the other hand to solar flattening, that is, with the fact that the figure of its rotation is an ellipsoid (Ref. 5). In accordance with the Jeans law the flattening of the inner layers should be greater than that of the surface layers. Consequently, the effect of differential rotation should be manifested more strongly in the inner, quite deep layers. However, this theory is contradictory to the law of rotation of active longitudes; as we have seen (see Chapter 2), their rotation is virtually rigid. Here, incidentally, we have an example of the use of data obtained from a study of solar activity for clarification of an extremely important general property of solar dynamics. In addition, in the extended convective zone the law of change of angular velocity with depth as established by Jeans and Rosseland cannot be satisfied, as we demonstrated in the preceding section. This forces us to reject the possibility of attributing the property of differential rotation to the flattening of a rotating star.

The most likely interpretation of the nature of the dependence of the angular velocity of rotation on astrographic latitude is the concept developed by Bjerknes, Rosseland, Krogdahl, Randers, and other authors (Refs. 66-69) that there is a baroclinic rotation of stars consisting of an ideal gas. One of the definitions of a baroclinic fluid, as is well known, is that it is one in which the surfaces of equal pressure and equal specific volume intersect, forming a system of temperature-pressure solenoids. The equation based on this property of baroclinic fluids was first derived by Rosseland (in vector form) and was then derived in coordinates by Klauder and later by Bjerknes (Refs. 70, 66).

We will give the derivation of this equation in coordinate form, taking into account its important value for the further exposition of the problem.

The equilibrium of a rotating mass of gas, if we neglect the influence of viscosity and assume that the z-axis coincides with the axis of rotation, is described by the following equations:

$$\begin{aligned} \frac{1}{\rho}\frac{\partial p}{\partial x} &= \frac{\partial V}{\partial x} + \omega^2 x, \\ \frac{1}{\rho}\frac{\partial p}{\partial y} &= \frac{\partial V}{\partial y} + \omega^2 y, \\ \frac{1}{\rho}\frac{\partial p}{\partial z} &= \frac{\partial V}{\partial z}, \end{aligned} \tag{2.145}$$

where ω is the angular velocity of rotation; p is pressure; ρ is density; and, V is gravity potential. Differentiating the first of these equations for y and the second for x, and substracting the one from the other, we obtain

$$\frac{\partial\left(\frac{1}{\rho}\right)}{\partial x}\cdot\frac{\partial p}{\partial y} - \frac{\partial\left(\frac{1}{\rho}\right)}{\partial y}\cdot\frac{\partial p}{\partial x} = 2\omega\left(y\frac{\partial\omega}{\partial x} - x\frac{\partial\omega}{\partial y}\right);$$

a similar transformation, also applied to the third equation, gives (combination of the second and third, and the third and the first)

$$\begin{aligned} \frac{\partial\left(\frac{1}{\rho}\right)}{\partial y}\cdot\frac{\partial p}{\partial z} - \frac{\partial\left(\frac{1}{\rho}\right)}{\partial z}\cdot\frac{\partial p}{\partial y} &= -2\omega y\frac{\partial\omega}{\partial z}, \\ \frac{\partial\left(\frac{1}{\rho}\right)}{\partial z}\cdot\frac{\partial p}{\partial x} - \frac{\partial\left(\frac{1}{\rho}\right)}{\partial x}\cdot\frac{\partial p}{\partial z} &= 2\omega x\frac{\partial\omega}{\partial z}. \end{aligned} \tag{2.146}$$

In order to convert to a vector representation of equation (2.146) it is necessary to multiply by the unit vectors $\bar{i}$, $\bar{j}$ and $\bar{k}$. Performing this multiplication and adding the products, we have:

$$\bar{i}\left(\frac{\partial\frac{1}{\rho}}{\partial y}\cdot\frac{\partial p}{\partial z}-\frac{\partial\frac{1}{\rho}}{\partial z}\cdot\frac{\partial p}{\partial y}\right)+\bar{j}\left(\frac{\partial\frac{1}{\rho}}{\partial z}\cdot\frac{\partial p}{\partial x}-\frac{\partial\frac{1}{\rho}}{\partial x}\frac{\partial p}{\partial z}\right)+$$
$$+\bar{k}\left(\frac{\partial\frac{1}{\rho}}{\partial x}\cdot\frac{\partial p}{\partial y}-\frac{\partial\frac{1}{\rho}}{\partial y}\cdot\frac{\partial p}{\partial x}\right)=2\frac{\partial\omega}{\partial z}(\omega x\bar{j}-\omega y\bar{i})+$$
$$+2\omega\bar{k}\left(y\frac{\partial\omega}{\partial x}-x\frac{\partial\omega}{\partial y}\right);$$

the components of the vector of linear velocity in the case of purely rotational motion are expanded relative to the x- and y- coordinate axes to the components $u = -\omega y$ and $v = \omega x$. By virtue of the condition of rotational symmetry, in addition,

$$y\frac{\partial\omega}{\partial x}=x\frac{\partial\omega}{\partial y}.$$

We simplify the above expression and obtain

$$2v\frac{\partial\omega}{\partial z}=\left[\operatorname{grad}\frac{1}{\rho}\operatorname{grad}p\right]. \tag{2.147}$$

This relation is of fundamental significance. It shows that the left-hand side of (2.146) will be equal to zero only when $\frac{\partial\omega}{\partial z}=0$ everywhere except at the axis of rotation, at which the linear velocity becomes equal to zero. In the other case noted, when the left-hand side is equal to zero, this will occur only when there is a coincidence of the surfaces of equal pressure and equal specific volume, i.e., in the case of a barotropic fluid.

However, in the case of a baroclinic fluid the angular velocity of rotation also is a function of the distance from the equatorial plane. When the right-hand term (2.146) equals zero, the angular velocity is a function only of the distance to the axis of rotation. When $\frac{\partial\omega}{\partial z}\neq 0$, the normals to the surfaces of both families of surfaces—of equal pressure and of equal specific volume—form a certain angle α. If $\frac{\partial\omega}{\partial z}>0$ the

flattening of the surfaces of equal pressure will be greater than that of the surfaces of specific volume; the reverse will be true if $\frac{\partial \omega}{\partial z} < 0$.

In the first case the temperature along the surface of equal pressure (isobaric), and the surface of specific volume (isosteric) increases from the pole to the equator; in the second case, the temperature decreases in this direction, i.e., the pole is somewhat hotter than the equator.

On the sun and stars it is characteristic that the second situation situation prevails, which also corresponds to a decrease of angular velocity in the direction of the poles. Thus, qualitatively the relation (2.147) successfully explains one of the principal properties of solar rotation.

Also of interest are conclusions based on a quantitative evaluation of relation (2.146). Taking the conditions at the equator of the sun and the mean values of the gradients for the entire depth when the value is $= 3 \cdot 10^{-6}$ (which, to be sure, is a certain exaggeration), we obtain the order of magnitude of the left-hand side of (2.146) equal to 10^{-11}. The order of magnitude of the right-hand side is $5 \cdot 10^{-7}$, and, therefore, the angle between the isobaric and isosteric surfaces is 4" (V. A. Krat gives 3", Ref. 71). However, in the layers close to the surface, for depths of 380 and 430 km below the base of the chromosphere ($\tau = 0.004$), it can be shown that the angle α will be approximately 10".

We should note the relationship between the result, on the one hand, and the first condition of stability of purely zonal circulation as determined by Helmholtz (Ref. 72); according to this condition the circulation is stable with an increase of potential temperature toward the pole of the corresponding hemisphere. It is obvious that circulation on the sun is not purely zonal: its characteristic feature is the presence of meridional flow, but nevertheless this circulation is stable (Ref. 62).

On the other hand, it is necessary to emphasize the close relationship between the relation (2.147) and the theorem on circulation of a baroclinic fluid—the Bjerknes theorem. As is well known, in contrast to the Kelvin-Helmholtz theorem, which is correct for a barotropic fluid and which has the form $\frac{dC}{dt} = 0$, according to the Bjerknes theorem

$$\frac{dC}{dt} = A + B + C', \qquad (2.148$$

where A is a term which takes into account the baroclinicity of the fluid; B is a term taking into account rotation; and, C is a term taking into account viscous forces. The first of these terms, i.e., A, can be represented in different ways, such as

$$A = N' - N'',$$

where N' is the number of positive, and N" is the number of negative isobaric-isosteric tubes (temperature-pressure solenoids) included within the particular circuit. We note that such tubes are considered positive if the expression is $-\int_l W dp = 1$ and negative if the expression is $-\int W dp = -1$; here W is the specific volume, p is pressure, l is the length of the elementary circuit, formed by parts of the traces of unit isobaric and isosteric surfaces (in this case, these surfaces are called unit surfaces because they are passed through one unit of calculation of specific volume and pressure respectively).

A different form of the first term of the right-hand side of (2.148) is obtained in a case when the circuit (i.e., the entire closed circuit) is formed by two isotherms with a temperature difference ΔT and two isobars with the pressures p_1 and p_2. In this case

$$A = R\Delta T \ln \frac{p_1}{p_2},$$

where R is the specific gas constant $R = \frac{\mathfrak{R}}{\mu}$.

Finally, the third form of the term A, approaching closest to the right-hand side of relation (2.147), is the expression of the baroclinic term A, which in the most general form is equal to $-\oint \frac{dp}{\rho}$. From this form we can derive a representation through the number of unit tubes and the special case when the circuits consist of isobars and isotherms, respectively (in so-called Silberstein form, Ref. 73). In this case, A is expressed as follows:

$$A = \sum_s \frac{1}{\rho^2} \frac{\partial p}{\partial n_1} \cdot \frac{\partial \rho}{\partial n_2} \sin \alpha;$$

here n_1 and n_2 are normals to the surfaces of equal pressure and equal density, and s is the circuit. If specific volume is introduced in place of density, the expression for A in Silberstein form will be identical to the expression in the right-hand side of (2.148).

The term B has the following form:

$$B = -2\omega \frac{dF}{dt};$$

here ω, as usual, is the angular velocity of rotation, and F denotes the area of the circuit in which circulation occurs. The change of this area with time can be explained graphically using the example of the trade winds (in certain respects, this example is close to what occurs on the sun).

We will define a circuit which encircles the earth along a circle of latitude near the northern boundary of the trade winds zone. Under the influence of a flow of air directed from north to south near the surface at these latitudes (in the northern hemisphere, of course), the circuit will shift in the direction of the earth's equator and will expand. This means that $\frac{dF}{dt} > 0$, but in accordance with the form of the term B this will denote the appearance of the component $\frac{dC}{dt} < 0$. This means the appearance of circulation along a particular circuit, in a clockwise direction as viewed from the north pole; in other words, an easterly wind component appears, and in place of a purely northerly trade wind, we obtain the usually observed NE Trades. There is basis to assume that something similar to the easterly component of motion (easterly wind) also exists in layers of the sun close to the photosphere. For example, this is manifested in the rather rapid displacement westward of forming groups (see Chapter 1). The displacement of the circuit along which the circulation occurs is caused in these cases by the presence of meridional currents, and this in turn is the result of a difference in temperatures, that is, in some way is associated with baroclinicity.

It is therefore possible to speak of a certain dependence between terms A and B of equation (2.148). It also is necessary to emphasize the need for a careful approach to statements of the following type which are sometimes encountered: a fluid, initially barotropic, in rotational motion behaves in relation to satisfaction of the circulation theorem as a baroclinic fluid, since the term B appears. Actually, the difference between B and zero means that the circuit is changing with time, that is, there exists some motion which usually is a result of baroclinicity.

In essence the term B takes into account the effect of the Coriolis force; development of the latter requires not only rotation, but also the presence of relative velocity. As already noted, the term

D takes into account the effect of viscous forces. In general, it is possible to interpret its value more broadly. We can write:

$$D = \int (X dx + Y dy + Z dz), \tag{2.149}$$

where X, Y and Z are the components of equivalent external forces, except pressure, gravity and Coriolis force. In this sense, in the case of a hydromagnetic fluid, the term D also includes the electromagnetic force, whose effect, taking into account the character of real magnetic fields, can in many respects resemble the effect of viscous forces.

Viscous forces can be represented as the integral of the product of the exchange coefficient and the Laplacian of velocity. Taking into account that the velocity of horizontal meridional currents is not great, as will be demonstrated soon, it is possible to apply the Guldberg-Mon model used frequently in dynamic meteorology, and assume that viscosity is proportional to velocity (much as in the study of convection under conditions of turbulent friction, resistance is assumed to be proportional to the square of velocity). In this case the term D will be expressed as follows (Ref. 14):

$$D = -nC.$$

The presence of viscosity is of extraordinarily great importance because it does not permit circulation to increase with time to infinity. It is true that there also are other factors which serve as an obstacle to such an increase, and the most important of these is the disappearance of the factors giving rise to circulation, that is, specifically the temperature difference, but viscosity is a factor which operates more rapidly in this respect. In actuality, if only factors A and B were operative and their effect was not mutually compensated, it would follow from the Bjerknes theorem that

$$C = A't,$$

where A' is a coefficient related to the combined effect of the factors A and B, i.e., the linear increase of circulation with time. However, in the case of a viscosity effect, instead of equation (2.148), we would have

$$\frac{dC}{dt} = A' - nC. \tag{2.150}$$

The solution of this linear equation with the same initial condition $C\big|_{t=0} = 0$ will be

$$C = \frac{A'}{n}(1 - e^{-nt}). \tag{2.151}$$

It can be seen from this relation that only for small t does circulation increase linearly with time, and then approach the constant value A'/n. The attainment of this state will correspond to equilibrium between the influence of baroclinicity and viscosity (Ref. 14). We will now attempt to make certain estimates of the characteristics of large-scale motions on the sun. In order to do so we require data on the meridional pressure and temperature gradients, and for this reason we will now take up the problem of the nature of the temperature and pressure differences between the high and low latitudes of the sun.

The history of this problem begins in 1924, the date of publication of the well-known Zeupel theorem. We note that this was a period when there were no concepts concerning even a thin convective zone, and an entire star, with the possible exception of the layers closest to the center, was considered to be in a state of radiative equilibrium. Investigating such a star and assuming that it rotates as a solid body, Zeupel found the law which in this case should determine the degree of generation of energy as a function of density (the Zeupel formula includes not only density, but also the angular velocity of rotation, the gravitational constant, and the number π). This law, however, does not agree with the real situation, because it follows from it that at a certain density value the generation of energy becomes negative, something without physical meaning. Eddington (Ref. 74) and Vogt (Ref. 75) independently concluded that the solution of this contradiction is possible by assuming the existence of large-scale meridional currents. If we use the center of a star (the sun) as the origin of coordinates, and take a cylindrical system in which the polar axis is directed along the axis of rotation of a star, in the case of solid-body rotation the equations of motion relative to the R- and z-axes will have the form

$$\left.\begin{aligned} \frac{\partial p}{\partial R} &= (g_R + R\Omega^2)\rho, \\ \frac{\partial p}{\partial z} &= g_z\rho, \end{aligned}\right\} \tag{2.152}$$

where Ω is angular velocity.

It is obvious that the forces represented in the left-hand sides of these equations (the dimensionality of force per unit mass, i.e., of acceleration, is obtained after division of both sides of each of the equations (2.152) by ρ) can be regarded as the components of the effective gravity potential ψ, that is, we can write

$$\left.\begin{aligned} -\frac{\partial \psi}{\partial R} &= g_R + R\Omega^2, \\ \frac{\partial \psi}{\partial z} &= g_z. \end{aligned}\right\} \tag{2.153}$$

As a mean for a certain surface, the levels of equation (2.153) are compatible with the condition of radiative equilibrium, but at individual points this condition is not satisfied if the Zeupel relation is not operative, which, as we have shown, is physically impracticable. Currents, therefore, should develop which tend to restore radiative equilibrium. Since R—a component of the effective gravity potential— is related to centrifugal force, as is clear from (2.152), an excess pressure should develop near the poles for equilibrium. It is associated directly with an excess of radiation at the poles, and causes a somewhat greater heating of the polar regions of a star than in its equatorial region. The degree of heating is associated directly with the acceleration of gravity, which changes with movement toward the pole from the equator (it increases).

This somewhat complex problem has been interpreted successfully by Milne (Ref. 76), and we wish to cite here the essence of his reasoning. The flux of radiant energy through the surface of a level at a certain point is given by the expression

$$F_n = -\frac{c}{\varkappa\rho}\frac{dp'}{dn}. \tag{2.154}$$

Here F_n is the flux in the direction of the normal n, p' is radiative pressure, and the remaining notations are the usual ones.

Milne introduces the expression

$$\lambda = \frac{1}{\varkappa}\cdot\frac{dp'}{p}, \tag{2.155}$$

here p is total pressure (gas + radiative).

By replacing $\frac{dp'}{x}$ by dpλ, we obtain

$$F_n = -\frac{c\lambda}{\rho}\cdot\frac{dp}{dn}.$$

The effective potential consists of two potentials—gravity and centrifugal. Therefore,

$$\psi = V + \frac{1}{2}R^2\omega^2,$$

where V is the gravity potential. Under conditions of relative equilibrium the force of the gradient of total pressure should be compensated by the combination of the gravity and centrifugal forces, that is, we should have

$$d\psi = \frac{1}{\rho}dp.$$

Substituting this expression into the formula for F_n, we have

$$F_n = -c\lambda \frac{d\psi}{dn} = c\lambda \frac{d\psi}{dR}$$

(since R and n coincide), but $\frac{d\psi}{dR}$ is gravity along a particular latitude, and F_n is proportional to T^4. Hence

$$T \sim \sqrt[4]{g_R}. \tag{2.156}$$

Thus, the pressure excess compensating the decrease of centrifugal force at the poles is an excess of radiative pressure, which combined with gas pressure gives an excess of total pressure.

The velocities of the meridional currents caused by the excess of radiation compensating the excess of gravity have been computed by many investigators: Sweet (Ref. 77), Schwarzschild (Ref. 3), Csada (Ref. 78) and others.

We will now present the essence of the reasoning of Eddington (Ref. 74). He considers not only the energy generated in a particular region of a star ϵ_1, but also the energy transported there by convection ϵ_2 (and, we might add, carried away by advection). It is obvious that all of his discussion concerns this value ϵ_2. Expressing the radiation flux as a function of the effective potential and its derivative along the normal, that is, the acceleration of gravity, Eddington writes the following rank equation

$$\text{rank } \epsilon_2 = \frac{f'(\psi)}{\rho} \cdot \text{rank } \left(\frac{d\psi}{dn}\right)^2;$$

here $f(\psi)$ is a function relating the radiant flux and gravity. Then the value $\pm q$ is introduced to characterize the limits of deviation of g^2 from its mean value, so that

$$\text{rank } g^2 = \pm q g^2, \tag{2.157}$$

the value q is dependent on the velocity of rotation of the star; for the sun, Eddington assumes q = 0.1, and finds the upper and lower limits for ϵ_2.

Expressing the energy transported by convection as the difference between the mean generation of energy for a particular star and its generation in a particular place (the corresponding notations are $\bar{\epsilon}_1$ and ϵ_1), Eddington obtains as the upper limit $\epsilon_2 \pm q\bar{\epsilon}_1$. In a unit volume the upper limit is $\epsilon_2 \pm q\bar{\epsilon}_1$. In a unit volume the upper extreme value ϵ_2 will be $\pm q\epsilon_1\rho$. Then Eddington considers, essentially, a thermodynamic equation of heat influx for the case of convective flow, and finds that this quantity can be represented as

$$c_v\rho T\left(\frac{\gamma-1}{\gamma}-\frac{1}{4}\right)\frac{dp}{\cdot p},$$

with

$$p=(\gamma-1)\,c_v\rho T,$$

where $\gamma = \frac{c_p}{c_v}$. Denoting the factor for dp by α, Eddington equates the convective currents, expressed thermodynamically, proceeding on the basis of considerations of gravitational and radiation properties. He obtains

$$\alpha dp = \pm q\bar{\epsilon}_1\rho. \tag{2.158}$$

Then using the equation of hydrostatic equilibrium, $dp = -g\rho dr$, Eddington finally obtains

$$v_r = \pm\frac{q\bar{\epsilon}_1}{\alpha g}, \tag{2.159}$$

since $v_r = \frac{dr}{dt}$; taking one side, he obtains $v_r = \alpha r$. Substituting here the above-cited numerical values (the value α used is the lower limit, $\alpha = 0.04$), Eddington finds that the upper limit for v_r will be $2 \cdot 10^{-4}$ $cm \cdot sec^{-1}$.

Later investigations by Sweet revealed that such a velocity is excessive. The difference $\bar{\epsilon}_1 - \epsilon$ is too high, but α, on the other hand, is too low. Sweet concluded that a velocity better corresponding to the present-day status of the problem is $v_r = 6 \cdot 10^{-9}$ $cm \cdot sec^{-1}$.

M. Schwarzschild (Ref. 3) determined a value of the order of 10^{-9} $cm \cdot sec^{-1}$. Such slow currents are representative of virtually all of the inner layers of a star. Matter rises along the axis of rotation to the polar regions and then flows down to the equator in the layers closer to the surface, moving in depth along the equatorial radius. Figure 33 is a diagram of this circulation. At such a velocity these currents (which have been named in honor of Eddington) would require more than 10^{12} years for the matter from the inner layers of the sun to be moved to its surface.

Since this time exceeds by a factor of more than 100 the age of the sun, these currents apparently are not capable of influencing phenomena associated with solar activity. This point of view has been basic since the time of the work of Eddington (1929). However, certain authors, including V. A. Krat (Ref. 9) and Wasiutynski (Ref. 11) have pointed out that it is necessary to distinguish Eddington currents from the circulation of a baroclinic fluid (Bjerknes circulation) associated with differential solar rotation; the latter circulation naturally cannot be caused by such a slow circulation as the Eddington-Sweet-Schwarzschild circulation model considered above. We therefore arrive at the logical inevitability of assuming the existence of two different physical models of circulation. However, we will first consider whether there are any factors capable of forcing Eddington currents to move with considerably greater velocity.

First, as Eddington himself noted, the velocity which he computed is vertical and the horizontal velocity can be greater; but as he points

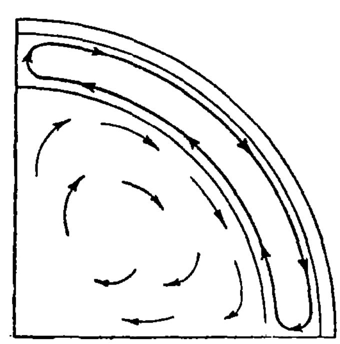


Figure 33. Diagram of circulation in the convective zone and in the inner zone of radiative equilibrium

out, the continuity equation does not allow an extremely great difference between these components. We can add to this that if the Eddington currents occupy a relatively insignificantly thick spherical layer, the difference between the vertical and horizontal velocities under certain circumstances could be rather substantial.

We have in mind here a model of a boundary layer occupying the entire celestial body (in dynamic meteorology, such a boundary layer is called "planetary", but in our case it naturally would be called "stellar"). As pointed out by N. Ye. Kochin, in such a boundary layer the vertical velocity is related to horizontal velocity as the thickness of the layer is related to the radius of the particular celestial body (Ref. 79). But this means of substantiating the possibility of an increase of horizontal velocity in comparison with vertical velocity first would require proof that there is circulation in the boundary layer, and then that this layer is thin in comparison with the solar radius.

The fact that the Eddington currents by no means occupy the entire thickness of the sun invalidates this means of proving the possibility of a considerable excess of horizontal over vertical velocity. However, it is possible to attempt, as Cowling has done in Ref. 8, to prove that the value α in formula (2.159) becomes very small near the convective zones. In accordance with (2.158), α becomes equal to zero when $\gamma = 4/3$. Cowling then correctly notes that in the convective zones themselves, there is no transport of energy by radiation, and as a result it becomes impossible to compute the vertical velocity there by use of the formulas derived by Eddington. However, according to (2.30) the value $\gamma = 4/3$ corresponds to $\nabla_{ad} = 0.25$.

Turning to the Vitense model we see that such a value of the logarithmic adiabatic gradient is attained when the logarithm of gas pressure is equal, first, to $\log p_g$ 5.2 (Biermann and his colleagues feel that this value of the adiabatic gradient corresponds to a somewhat larger logarithm of gas pressure, equal to approximately 5.4); and, second, at a value $\log p_g = 7.7\text{-}8.0$. In essence, both the upper and lower values lie in the convective zone, but outside the zone the adiabatic gradient very rapidly attains a value of 0.4. Therefore, for all practical purposes it is impossible to use this means to prove a significant increase in the velocity of the currents.

In comparing the Eddington currents and the circulation of a baroclinic fluid, it is impossible to overlook the fact that in all the above discussions the properties of the latter were taken into account inadequately.

The dependence of a temperature increment in the high latitudes of a star on the increment of acceleration of the gravity there was derived by Milne on the basis of the circumstance that an increase of temperature, as a result of a somewhat greater radiation flux, leads to an increase of radiative pressure, and as a result there is also an increase in total pressure. However, it is well known that in the case of the sun, radiative pressure is but a small part of the total pressure. It is possible to conceive of the following model: the temperature increase in the high latitudes of a star, caused by certain factors (nonuniform distribution of sources of heating), leads to an increase of gas pressure along an isosteric surface in accordance with the equation of state.

It must be noted, however, that in determining the velocity of meridional currents such a possibility is taken into account indirectly, since the results are obtained from quite general methods of consideration, based on the energy characteristics of the phenomenon. A typical example of such an approach is the interpretation of the problem by Schwarzschild. Although he considers the conditions of radiative equilibrium, in principle his method can be used under other conditions, after appropriate modifications.

The source of the energy transported by circulation is the divergence of the radiation flux. Schwarzschild equates this divergence to losses of energy as a result of circulation consisting of two terms: the work done on the surrounding medium, and the divergence of the convective transport of heat. As already mentioned, this computation gives a small value of vertical velocity. However, even the most elementary consideration of large-scale motion on the sun as currents of a baroclinic fluid leads to completely different values of horizontal velocity. In this connection we will mention a study (Ref. 80) we made in 1954.

In our investigation it was demonstrated that if we proceed on the basis of the zonal gas pressure gradients, it is possible to obtain a velocity of meridional currents capable of explaining the duration of the 11-year solar cycle on the basis of the time required for movement of the spot-forming zone from the high to the low heliographic latitudes. The author considered motion applying to the class of geostrophic phenomena (in this case it would be correct to say heliostrophic). Prior to this investigation another study was made in which it was possible to demonstrate that as a result of the conservation of angular momentum the circulation along meridians should be differential. This study will be discussed in greater detail below, but at this point it is important that in the simplest case of inertial motions, the displacement along a meridian can occur only under the influence of a zonal gradient; with respect to its numerical value, it was assumed equal to its meridional gradient.

We will now give the reasoning that guided us in the study cited, taking into account, however, the necessity for a certain improvement of those computations which were made in Ref. 80. Most important is that the depth beneath the photosphere at which the zone of radiative equilibrium begins, assumed in Ref. 80 to lie at a distance of $3\text{-}4 \cdot 10^9$ cm from the surface of the layers, now is considered equal to $9.3 \cdot 10^9$ cm, in accordance with the Motz-Epstein model. This naturally changed the assumed value of density and molecular weight.

Somewhat different concepts were later developed relative to the possibility of taking into account longitudinal relative velocity. We therefore feel it is desirable to successively discuss the concepts developed in Ref. 80 with those changes introduced in connection with the study of this problem during the last eight years.

The equation giving the angular velocity $\dot{\varphi}$ of the meridional wind has the form

$$2r(\lambda+\omega)\sin\varphi\cdot\dot{\varphi}=\frac{\partial p}{r\rho\cos\varphi d\lambda}, \qquad (2.160)$$

where φ is latitude; λ is longitude; ω is the angular velocity of solid-body rotation; $\dot{\lambda}$ is the correction to angular velocity necessitated by proper motions or differential rotation; p is pressure; r is the radius vector; and, ρ is density. Using the notations $\dot{\lambda}+\omega=\omega_1$, we have:

$$\dot{\varphi}=\frac{\partial p}{r^2\rho\omega_1\sin 2\varphi d\lambda}. \qquad (2.161)$$

The pressure drop along the surface of equal density is related to the temperature drop as follows

$$\Delta p=\frac{\Re\rho}{\mu}\Delta T. \qquad (2.162)$$

In order to find T we write the relation (2.156) in the form

$$T_\varphi=K\sqrt[4]{g_\varphi}, \qquad (2.163)$$

where K is the proportionality factor. Putting this relation into logarithmic form and differentiating, and then substituting the result into (2.162) we obtain

$$\Delta p=\frac{\Re\rho}{\mu}\cdot\frac{1}{4}T_\varphi\frac{\Delta g_\varphi}{g_\varphi}. \qquad (2.164)$$

We also should change ω, taking into account that a rather considerable depth is being considered. However, as already pointed out, within the convective zone the elementary Jeans law is not satisfied. In addition, with increasing depth there also is an increase of g_φ, which enters into the denominator. Therefore, replacing ω and g_φ by their values near the surface, we compensate their mutual change.

It is still necessary to find g_φ and $\Delta\lambda$. The first is determined from the relation

$$g_\varphi = g_\pi - 2\omega^2 r \cos^2 \varphi, \tag{2.165}$$

where g_π is the acceleration of gravity at the pole. Differentiating this expression for φ with r constant, and converting to finite differences, we have

$$\Delta g_\varphi = +2\,\omega^2 r \sin 2\varphi \Delta\varphi. \tag{2.166}$$

Substituting (2.166) into (2.164), (2.164) into (2.162), and (2.162) into (2.161), and taking into account that for the equality of gradients when there is an equality of dropoffs, we should have $\Delta\varphi = \Delta\lambda$, we find after substitution that $\dot{\varphi} = 4.7 \cdot 10^{-8}$. The equality of $\Delta\varphi$ and $\Delta\lambda$ will be exact when $\varphi = 45^\circ$, but since $\bar{\varphi} \approx 15^\circ$, then $\Delta\varphi = 1/2\,\Delta\lambda$. With an accuracy to this factor it can be assumed that after shortening by $\sin 2\varphi, \Delta\varphi = \Delta\lambda$. If it is taken into account that movement along a meridian in a solar cycle in abstract terms averages 0.309, it becomes clear that at such a velocity such movement will require approximately 2.5 months. However, if we substitute into the formula for $\bar{\dot{\varphi}}$ the conditions for the photosphere (the principal difference will of course be temperature) it becomes clear that the passage along the same meridional arc will require about 33 years.

It is obvious that with the simplifications made in this computation it is difficult to expect great agreement. The value obtained for the photosphere exceeds by only a factor of 3 the true length of the 11-year cycle, and this cannot be regarded as other than a success in such elementary calculations. In these computations motion was considered for the case of equivalent forces, and viscosity was neglected completely in the computations. In general, the problem of taking viscosity into account is by no means simple.

Even those velocities of meridional movement which were just derived do not exceed the mean wind velocity in the circulation of the

earth's atmosphere (of the order of 10^3 cm·sec^{-1}). The coefficient of kinematic turbulent viscosity is equal in order of magnitude to the product of velocity and the scale of motion. The latter on the sun is two orders of magnitude greater than on the earth; but from application of the theory of similarity to atmospheric motions it is known that the larger the scale of motion, the lesser is the role of viscosity.

It is therefore by no means precluded that the value of viscosity in solar meridional circulation does not greatly exceed this value in the general circulation of the earth's atmosphere. This circumstance can be used as a certain basis for computations of the type just made, i.e., computations in which viscosity is not taken into account. This of course does not exclude the necessity for taking into account viscous forces in a whole series of cases, such as in the investigation of the theorem of circulation, which was discussed above.

The computations made, despite their primitive character, convincingly demonstrate that there is no correlation between the vertical velocity of Eddington currents and the horizontal velocity of meridional currents. We have seen that in the case of satisfaction of the continuity equation, such a situation could not occur, as already noted by Eddington. Therefore, a natural conclusion appears evident: under the conditions of general circulation on the sun, the continuity equation may not be satisfied.[1] This means that there is a subphotospheric spherical layer constituting a discontinuity. It is possible that there are several such layers. Such a situation indicates the phenomenon of circulatory instability, which together with gravitational and rotational instability, constitutes an important property capable of facilitating the development of nonstationary phenomena, i.e., phenomena of importance for solar activity.

We will now turn to the problem of differentiation of solar circulation. Meridional flows on a rotating sphere cannot be maintained under ordinary conditions due to the appearance of a zonal component of velocity associated with the conservation of angular momentum. This is the situation on the sun. We considered this elementary problem in 1953 (Ref. 82). We will repeat the basic reasoning here, and by anology with the earlier study will present computations based on new data concerning these problems.

---

[1]It is true that the studies of Kippenhahn, which will be discussed below, give an appreciably greater vertical velocity for Eddington currents and make it possible to satisfy the continuity equation.

It is known that a meridionally oriented deflecting force, operating against the force of the pressure gradient, is equal to:

$$A = 2\omega u \sin\varphi, \tag{2.167}$$

where, once again, ω is the angular velocity of solar rotation; φ is latitude; and, u is the zonal component of velocity, developing at the time of the meridional displacement of a point as a result of conservation of angular momentum. If the point moves from the latitude φ to the latitude φ', this component will be (Ref. 82):

$$u = r\omega \frac{\cos^2\varphi' - \cos^2\varphi}{\cos\varphi'}, \tag{2.168}$$

where r is the radius of the celestial body. Substituting (2.168) into (2.167), we have

$$A = 2r\omega^2 \frac{\cos^2\varphi' - \cos^2\varphi}{\cos\varphi'} \sin\varphi; \tag{2.169}$$

here φ is the mean value of latitude in the interval (φ', φ).

In order for the meridional currents to move without obstacle, the force of the corresponding pressure gradient, that is, the quotient from the division of the gradient by density, should be greater than the deflecting force given by formula (2.169), that is

$$\frac{\Delta p}{\rho r \Delta\varphi} > 2r\omega^2 \frac{\cos^2\varphi' - \cos^2\varphi}{\cos\varphi'} \sin\bar{\varphi}. \tag{2.170}$$

The meridional pressure gradient is determined from the temperature drop by use of formula (2.162); however, the temperature drop for the zone of radiative equilibrium, including the photosphere, is determined from (2.163). By making the appropriate computations we obtain for the beginning of the inner zone of radiative equilibrium (the first outer point in the Motz-Epstein model), a gradient with the force $G = 4.7 \cdot 10^{-4}$ $cm \cdot sec^{-2}$, and for the photosphere $G = 1.7 \cdot 10^{-4}$. The decrease of the gradient is associated primarily with an increase of molecular weight.

In the photosphere, μ was assumed to equal 1.3, and in the inner zone of radiative equilibrium, 0.52. Both cited values of the force of the pressure gradient were obtained for a temperature drop between the pole and equator of 0.3° (Ref. 81). For the purpose of a comparison with the Coriolis force, which is an obstacle to the motion of meridional currents, we cite Table 23 (taken from our earlier study, Ref. 81).

Table 23

φ	φ'	u, $cm \cdot sec^{-1}$	A, $cm \cdot sec^{-2}$
90°	80°	361	$2.16 \cdot 10^{-1}$
80	70	530	3.07
70	60	555	3.02
60	50	530	2.60
50	40	473	2.00
40	30	392	1.35
30	20	298	$7.6 \cdot 10^{-2}$
20	10	181	2.8
10	0	62	$3 \cdot 10^{-3}$

This table gives the zonal velocities and deflecting forces developing with a transition from the heliographic latitude cited in the first column to the latitude in the second column.

It follows from the table that the Coriolis force developing with meridional movement on the sun is capable at all latitudes of preventing this movement, if the force of the gradient has the order of magnitude that we have determined. This means that in the case of the temperature drop occurring under conditions of radiative equilibrium, that is, associated with the change of gravity with heliographic latitude, a longitudinal differentiation of circulation is inevitable. We will next discuss briefly other variants of the solution of this problem; however, we will now consider the significance of such differentiation.

By analogy with the simplest model of circulation of the earth's atmosphere (incidentally, in many respects outdated), this means that general circulation on the sun consists of a number of regions alternately having cyclonic and anticyclonic rotation. An analog of the first actually is observed in active regions, as was noted in 1941 by R. S. Gnevyshev (Ref. 83). However, no large-scale analog of anticyclonic rotation is observed on the sun, and the latter can be situated at another level, which is relatively inaccessible to observations. In the differentiation of circulation (an elementary model of such differentiation has been shown in Figure 34), we can see that there should be a longitudinal alternation of meridional currents directed toward the equator and toward the pole in a particular hemisphere.

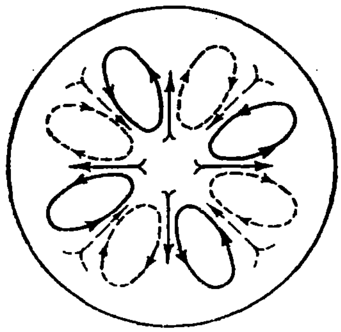


Figure 34. Model of differentiation of circulation

The actual pattern of motion of solar active formations along meridians is such that at least within the confines of latitudes 0-15°, sunspots reveal a tendency to move in the direction of the equator. However, it must be noted that formations in the higher layers of the solar atmosphere, specifically filaments, at a particular stage of development of an active region have a tendency to move in the direction of the pole (see Chapter 1).

It is true, as shown by the investigations of d'Azambuja (Ref. 8), that there are no significant regular longitudinal displacements of filaments relative to spots: filaments usually are situated at the same longitudes as spot groups. Therefore, whereas the model in Figure 34 is theoretically applicable to solar circulation, the alternation of cyclonic and anticyclonic centers (it would be better to say polar and equatorial currents) does not occur longitudinally, but by height: equatorial currents predominate in the photosphere and polar currents in the chromosphere. It should be noted, however, that due to lesser density and other factors, electromagnetic forces can play a more significant role in the chromosphere.

In the convective zone the problem of the temperature drop between the high and low solar latitudes naturally should be considered differently than in the zone of radiative equilibrium. We first will mention certain empirical data. Already, over a period of many years, individual investigators have attempted to clarify the problem of the temperature drop observed on the sun between the pole and equator by proceeding directly on the basis of observations. In the review published by Beckers (Ref. 84) somewhat recently the author notes three methods for investigation of this difference: using the color index; on the basis of the difference in radiation; and, interpretation of the difference in intensity of absorption lines.

It is easy to see that the first two methods can give a drop at the level of the photosphere, whereas the third indicates a drop at the level of the reversing layer. The conclusion in general is essentially that the drops are lower on the basis of photospheric data and higher for the reversing layer. In the case of the photosphere the most reliable results are those of Minnaert. He found that the solar pole on the average is 15° hotter than the equator. For the reversing layer Beckers obtained a drop of 60° (in the same sense). It was established further that the value of this drop is dependent on the phase of the 11-year solar cycle, and near the epoch of the minimum the equator is hotter than the poles. However, in general the temperature minimum is not situated at the equator, but at a latitude of 45°.

If Minnaert's data are accepted, it becomes clear that the force of the gradient will be $8.7 \cdot 10^{-3}$ cm$\cdot$sec^{-2}. This means that in a zone approximately 15° from the equator, the meridional currents can propagate without any obstacle, which is in good agreement with the results of investigations of the movements of sunspot groups themselves in a meridional direction (see Chapter 1). However, we are not inclined to draw far-reaching conclusions from the presence of a singular "trades" circulation on the sun, first because the temperature drop between the pole and equator was determined very unreliably by Minnaert, and second because the theory developed above is very primitive. It should be noted that Tuominen (Ref. 85), Schoenberg (Ref. 21), and then Plaskett (Ref. 86) came close to consideration of similar problems, but without particular success.

The differentiation of circulation, that is, the development of zonal gradients ensuring meridional movements under conditions of satisfaction of the "heliostrophic balance" is, however, only one of the possibilities of overcoming the obstacle confronting meridional circulation as a result of the law of conservation of angular momentum.

There is another possibility, which in essence is that since the sun has no solid surface, the distance from the axis of rotation can remain constant during the movement of a point along a meridian due to sinking into deeper layers (when moving toward the equator), or rising into higher layers (when moving toward the pole). In addition, since the effect of the deflecting force under certain circumstances nevertheless can create a zonal component of velocity, the path of the particle in the meridional flow will constitute a spiral. However, quantitatively such a situation will not correspond to the real situation, since the vertical velocities developing under conditions of conservation of angular momentum will be several orders of magnitude lower than the zonal velocities developing by virtue of these same factors. Thus, this variant must be rejected.

There is still another possibility, consistent with the idea of Prandtl (Ref. 87). Prandtl assumes that at a very high viscosity the meridional currents remain such, despite the effect of the deflecting force. A qualitative interpretation is given best on the basis of the Jeffreys wind classification (Ref. 73). In the case considered, in the equations of horizontal motion which we write as follows:

$$\left.\begin{aligned} \frac{du}{dt} - 2\omega v \sin\varphi &= -\frac{1}{\rho}\frac{\partial p}{\partial x} + K\frac{\partial^2 u}{\partial z^2}, \\ \frac{dv}{dt} + 2\omega v \sin\varphi &= \frac{1}{\rho}\frac{\partial p}{\partial x} + K\frac{\partial^2 v}{\partial z^2}, \end{aligned}\right\} \qquad (2.171)$$

the terms representing frictional forces exceed both terms with the Coriolis force and terms taking acceleration into account, that is

$$K\frac{\partial^2 u}{\partial z^2} \gg \frac{du}{dt}; \quad K\frac{\partial^2 v}{\partial z^2} \gg \frac{dv}{\partial t}; \quad K\frac{\partial^2 u}{\partial z^2} \gg r\omega v \sin\varphi;$$

then equations (2.171) reduce to:

$$\left.\begin{aligned} K\frac{\partial^2 u}{\partial z^2} &= \frac{1}{\rho}\frac{\partial p}{\partial x}, \\ K\frac{\partial^2 v}{\partial z^2} &= \frac{1}{\rho}\frac{\partial p}{\partial y}. \end{aligned}\right\} \qquad (2.172)$$

In this case the wind blows in the direction of the gradient, and the magnitude of its velocity is regulated by friction. However, it is known from investigation of the earth's atmosphere that this class of winds, called antitryptic, is characteristic of small-scale phenomena (breezes, certain kinds of mountain-valley winds, involving a relatively small part of the surface), and does not correspond to phenomena with a scale commensurable with the general circulation of the earth's atmosphere. As a result, there is no basis for assuming that for some reason the large-scale circulation in the solar atmosphere will be more dependent on friction than on the Coriolis force.

We have already discussed the problem of the role of friction in solar circulation, and we concluded that the increase of turbulent viscosity on the sun in comparison with the conditions prevailing on the earth is dictated by the greater scale of solar movements, and should be compensated by a decrease in the importance of viscosity with an increase in the scale of the movement (in accordance with the conclusions of the theory of similarity). In summary, it can be said that currently the most acceptable explanation for the presence of stable meridional currents, in our opinion, is the presence of a zonal pressure gradient on the sun.

To be sure, in theory it is possible to postulate still another explanation of the conservation of meridional currents based on magnetohydrodynamic considerations. If in actuality there was a quite strong poloidal field, that is, if the lines of force were in meridional planes, the meridional currents would move in singular magnetic "guides", and could not deviate to any side. However, taking into account the weakness of the sun's general magnetic field, it appears improbable that the limitation of currents by the magnetic field could be of appreciable significance. It is more likely that in the presence of a field, additional viscosity develops, and this of course behaves differently in relation to the scale of movement than hydrodynamic viscosity, but is nevertheless incapable of exerting a quite considerable resistance to the deflecting force, and with it, resistance to differentiation of circulation.

We will now return to the velocity of meridional currents. A number of studies made in recent years by Kippenhahn (Refs. 88-91) deal with this problem, specifically, problems involved in the rotation of stars. Kippenhahn begins, to use his expression, with the theory of a "zero order", and then proceeds to more perfect theoretical models. The "zero order" theory is applicable to very slowly rotating stars, and in practice should be replaced by a model of higher approximations. The "first order" theory takes into account the role of meridional currents. Developing these concepts, Kippenhahn, in collaboration with Baker (Refs. 90, 91), published an article devoted in particular to meridional circulation in the case of rotation of a star different from rigid.

These two last studies (Kippenhahn and Baker) were devoted in particular to the problems considered here; Kippenhahn establishes that for the zone of radiative equilibrium, the vertical velocity of circulation is determined using the formula:

$$c_p\rho\left(|\mathrm{grad}\,T| - |\mathrm{grad}_{\mathrm{ad}}\,T|\right)v_r \approx \frac{L_r}{M_r}\rho x, \tag{2.173}$$

where c_p and ρ have the usual values. The gradients are taken in ordinary rather than in logarithmic form; L_r is luminosity as a function of distance to the center; M_r is the same for mass; and, x is given by the equation

$$x = \frac{\omega^2}{4\pi G\rho}, \tag{2.174}$$

where G is the gravitational constant. This, as noted by Kippenhahn, corresponds closely to the expression derived by M. Schwarzschild. But Kippenhahn assumes that in the right-hand side of (2.173) density should

be replaced by mean density $\bar{\rho}$, which obviously leads to a considerable increase of vertical velocity. For example, in the case of a type B star, the velocity near the surface is 6 orders of magnitude higher than in the inner regions of the star.

However, if we compute the vertical velocity of circulation by use of formula (2.173) for the outermost points of the inner zone of radiative equilibrium of the sun, we obtain $v_r = 4.5 \cdot 10^{-7}$ cm/sec. Here it is rational to use the Weyman model, although with an accuracy to an order of magnitude the same values also are derived from the Motz-Epstein model. This is even lower than Eddington's initial values, and only slightly exceeds the results obtained by Sweet. For the photosphere, the vertical velocity computed using this same formula is $6 \cdot 10^{-3}$ cm/sec. It can be assumed that the horizontal velocity will be greater by an order of magnitude or somewhat more, but there is no basis to assume that it will exceed the vertical velocity by almost 5 orders of magnitude, which is necessary for explaining the velocity of the intrinsic movements of sunspots along the meridian.

If the photosphere could be considered as a boundary layer, the vertical velocity in it could be related to horizontal velocity as the thickness of the photosphere is related to the solar radius (Ref. 79), that is, the horizontal velocity would be of the order of 10 cm/sec, which is an order of magnitude lower than the meridional velocity both of the intrinsic movements of spots and the movement of the spot-forming zone from the high to the low latitudes. There is, however, no basis for such a consideration, since according to Biermann the meridional velocity in the convective zone of the sun is such that the currents travel the distance from the high to the low latitudes in approximately 10 years. This gives a linear velocity of the order of 10^2 cm/sec.

The break in the velocity at the upper boundary of the convective zone is only one order of magnitude, and in fact cannot be called a real break in that the upper boundary of the convective zone is not expressed too clearly. It is therefore difficult to say that the attachment conditions are satisfied at the upper boundary of the convective zone, and assume the photosphere to be a "stellar" boundary layer. Thus, in order to explain phenomena on the sun, an alternative must be found to the Kippenhahn theory. Biermann, who obtained the value $v_m = 1$ m/sec for the velocity of currents in the convective zone, has supplied a satisfactory result.

Comparing this result with that obtained from computations based on use of formula (2.173) for the velocity at the outer points of the inner

zone of radiative equilibrium, we see that the break in horizontal velocities at the base of the convective zone is extremely significant. The inner zone of the radiative equilibrium can be considered in this sense to be fixed relative to the convective zone. If we agree with Vitense concerning the rather sharp transition from conditions of convection to conditions of radiative equilibrium at the lower boundary of the convective zone, it can be stated that the Kippenhahn theory, in combination with Biermann's computations, has introduced no appreciable changes to the concepts that we advanced in 1958 concerning the possibility of considering the convective zone as a "stellar" boundary layer (Ref. 19). The certain contradiction here can be viewed only from that elementary theory based on motion in the presence of equivalent forces, and under the influence of a zonal pressure gradient. This gives a value for meridional velocity at the level of the upper layers of the inner zone of radiative equilibrium, which exceeds by more than an order of magnitude the value computed by Biermann for the convective zone.

It could be postulated that the zonal gradients directly below the lower boundary of the convective zone are considerably less than the meridional gradients, that is, the conditions prevailing there are close to purely zonal circulation, although, to be sure, such conditions are not attained because meridional currents do not cease to exist. However, such a concept must agree with the fact that the law of rotation of active longitudes is a solid-body law. It is scarcely possible to count on viscosity playing a role in these layers because turbulent viscosity apparently is without appreciable significance there, and radiative and gas kinetic viscosity are incapable of decreasing so greatly the velocity of the meridional flow. Computations show that the term which takes viscosity into account will be many orders of magnitude smaller that the term associated with Coriolis force.

It would be possible to solve this problem by assuming that the active longitudes are associated with the lower layers of the convective zone where there already is virtually solid-body rotation. In this connection we must return to the problem of the distribution of differential rotation with depth.

Distribution of Differential Rotation With Depth

We already have pointed out that according to the Bjerknes-Rosseland formula, the angle of inclination between the isobaric and isosteric surfaces for the photosphere is in order of magnitude about 10". We will now discuss in greater detail the possibility of a change with depth beneath the photosphere of both the product of the moduli of the pressure and specific volume gradients, and the angle of reciprocal inclination of the corresponding isosurfaces. Expanding the right-hand side of (2.147), we can write

$$2\omega l \frac{\partial \omega}{\partial z} = \operatorname{grad} p \operatorname{grad} \frac{1}{\rho} \sin \chi. \tag{2.175}$$

Thus, a change of angular velocity with distance from the equatorial plane is dependent both on the moduli of the pressure and specific volume gradients, and on the angle of reciprocal inclination of the isobaric and isosteric surfaces. We will consider the second factor first. The angle of reciprocal inclination of these surfaces determines (all other conditions being equal) the drops of temperature and pressure along the surface of equal density (or equal specific volume); thus, this angle is associated with the distribution of energy sources by heliographic latitude; the greater the angle, the greater is the temperature and pressure difference between the high and low latitudes of the sun (star).

It is obvious that there should be a physical cause of such a difference, that is, from the hydrodynamic point of view, baroclinicity should be the physical cause. This factor, as noted by V. A. Krat (Ref. 9), is associated with the nonuniform distribution of energy sources. Due to some factor the pole is heated more strongly than the equator of a self-radiating gaseous celestial body. Methodologically it would be incorrect to attribute this to the circumstance that otherwise the star simply would not be stable. Such a mathematical-mechanical approach in the long run could lead us to teleological concepts.

The simplest physical explanation of the more considerable heating of the poles in comparison with the equator of the sun is flattening as a result of rotation. As pointed out by S. B. Pikel'ner (Ref. 10), "the polar parts of a flattened spheroid are closer to the center, where energy is released, i.e., the optical thickness along the polar radius is less than along the equatorial radius. The energy flux in the polar regions, therefore, is greater than in the equatorial regions, and the polar regions are therefore hotter" (p. 238).

This simple explanation, however, meets with the same difficulty that we encountered in an attempt to explain differential rotation on the basis of flattening; if this were true, the active longitudes would not reveal rigid rotation since differential rotation at great depths would be expressed more strongly, rather than more weakly, than in the layers closer to the surface. Since this is not the case, it is necessary to find another explanation of the temperature and pressure drop between the high and low latitudes of the sun. A certain role also can be played by the general magnetic field on the condition that it is an internal field, and one induced from the outside (Ref. 92).

Obviously, the propagation of an Alfven wave (and possibly waves of a different type) along such a field and perpendicular to it, occurs

differently and with different Joule and other losses. Unfortunately, this problem is extremely complex, and we are not able to develop this point of view.

With respect to the first factor, that is, the moduli of the pressure and specific volume gradients, their product decreases with depth, and since there is no basis for considering the angle between these vectors to be greatly changeable with increasing depth within the sun, this should mean a decrease in the degree of differential rotation with depth. We already have developed such concepts in Ref. 93. However, the table accompanying that article contained an error in computations; in addition, it is rational to make computations separately for the photosphere, separately for the convective zone,and separately for the very deep layers (the inner zone of radiative equilibrium).

In the latter case it is possible to proceed on the basis of the distribution of pressure, temperature and density with depth, which is indicated by the Motz-Epstein model, because it is that model which is used in our book. In the computations, the angle X was assumed equal to 10", which should be close to the actual value. As already mentioned, this angle cannot increase too greatly with depth because otherwise in such layers there would be a great temperature drop between the high and low latitudes, which would not remain undetected even in the surface layers (the degree of differential rotation also could not increase if an increase of the angle was compensated by a decrease of the moduli of the gradients). Particular attention must be given to the problem of differential rotation in the convective zone.

It was pointed out by Biermann in 1956 (Ref. 94) that the Bjerknes-Rosseland theorem in its initial form is unsuited for the convective zone. Biermann formulated the problem of deriving a theorem for the conditions applying in an extensive convective zone similar to the abovementioned theorem and applicable to the zone of radiative equilibrium. Biermann's basic premises can be reduced to the following two theories: First, turbulent exchange in the convective zone is not isotropic, since in the direction of gravity there should be an appreciable predominance, in particular, of transport of angular momentum. At the same time, Biermann notes a similar situation in the earth's atmosphere, which we feel is inapt. It is well known that the coefficient of vertical turbulent exchange in the earth's atmosphere is several orders of magnitude lower than the coefficient of horizontal exchange, also called macroturbulent exchange. This is true not only for small ascending currents, but also for exchange in such relatively large formations as cumulus clouds, which, it might be mentioned, sometimes occupy the entire thickness of the troposphere. The problem involves the characteristic dimensions of the corresponding vortices. Tropospheric cyclones and anticyclones are in size incomparably greater than cumulus clouds (with respect to horizontal extent, and with the same order of scale

vertically). Consequently, in discussing macroturbulent exchange on the sun it must be remembered what macroturbulent formations are responsible for it. We assume that active regions, to some degree, outline these enormous elements of solar macroturbulent exchange, although the absence of visible activity still is not indisputable evidence of the absence of such elements. In support of the Biermann concept, it is possible to advance the concept of great vertical velocity of convection in comparison with the velocity of circulation. In actuality, under terrestrial conditions the vertical velocities in the atmosphere are considerably less than the horizontal velocities, except for such rare formations as waterspouts, tornadoes, etc. However, it must be remembered that the mixing length is considerably less than the thickness of the extensive convective zone, so that the coefficient of vertical exchange in the convective zone apparently nevertheless is not greater than the coefficient of horizontal exchange. It, therefore, must be assumed that the right-hand side of the formula derived by Biermann is as indicated below

$$\frac{d \ln \omega}{d \ln r} \approx \frac{-2A_2}{A_1 + A_2}, \tag{2.176}$$

where A_1 is the isotropic and A_2 is the "monotropic", that is, vertical component of the exchange coefficient; as a result of the fact that $A_2 \approx A_1$, the right-hand side of formula (2.176) should be a value of the order of unity, which agrees with the result obtained above in the solution of the Jeans equation for the convective zone.

Biermann proceeds secondarily from the fact that the equation for horizontal motion in the convective zone does not have the form in which formula (2.147) would be correct, that is, the Bjerknes-Rosseland theorem in its usual form. This point of view is undoubtedly correct, since such a result follows directly from the Navier-Stokes equations if they are used in place of (2.145). The same conclusion was drawn long ago by Randers in Ref. 69. The problem has been investigated by many authors and in particular detail by Csada in Ref. 17.

The problem is as follows. The Bjerknes-Rosseland theorem, which we derived earlier in this chapter, can be obtained in vector form as well as in coordinates. The equation of equilibrium for a rotating ideal fluid can be written as follows:

$$[\bar{\omega}[\bar{\omega}\bar{r}]] = \nabla\Phi - \frac{1}{\rho}\nabla p, \tag{2.177}$$

where Φ is the gravity potential. Taking the curl from both sides of (2.177) and taking into account the incompressibility of the fluid, and

also that is dependent only on the z coordinate, after transformations it is possible to obtain the Bjerknes-Rosseland formula (2.146). However, in the case of a viscous fluid the equation (2.147) will include a term (and when incompressibility is taken into account—two terms) characterizing viscous forces. After removal of the curl the right-hand side no longer will be the same as the right-hand side of (2.177) but it will include the Laplacian of the curl of velocity with the coefficient of viscosity as a factor. This same expression is given by Biermann, whose formula has the form $\mathrm{curl}[\bar{\omega}[\bar{\omega}\bar{r}]] = \mathrm{curl}\left(\frac{1}{\rho}\Delta p\right) +$ terms associated only with v_m; (2.178) here v_m is meridional velocity.

It is physically obvious that the terms which take viscosity into account should be an obstacle to differential rotation, that is, they should be subtracted from the value grad p·grad $1/\rho$. Therefore, in the convective zone, the product of the linear velocity of rotation and the derivative of angular velocity along the distance from the equatorial plane is not compensated for by the product of the pressure and specific volume gradients (Kippenhahn, Ref. 89). The values $\frac{\partial\omega}{\partial z}$ for the convective zone, derived from the ordinary Bjerknes-Rosseland formula, therefore, are too high; they should be regarded as upper limits. Figure 35 shows log $\frac{\partial\omega}{\partial z}$ as a function of depth beneath the photosphere. It can be seen from Figure 35 how rapid is the decrease of the dependence of differential rotation on distance to the equatorial plane with increasing depth. It should be noted that in the region of the inner zone of radiative equilibrium the values $\frac{\partial\omega}{\partial z}$ should be still less since the computations were made by us on the assumption of a constancy of ω with depth, although in actuality in the region of radiative equilibrium this value should change in accordance with the Jeans law. Taking into account that the derived values $\frac{\partial\omega}{\partial z}$ for the convective zone are upper limits, it is necessary to regard rotation there as almost barotropic.

Thus, if we do not propose the physically almost unlikely hypothesis that there is a considerable increase of the angle of reciprocal inclination of isobaric and isosteric surfaces with depth, it must be concluded that differential solar rotation is a property of layers close to the sun's surface.

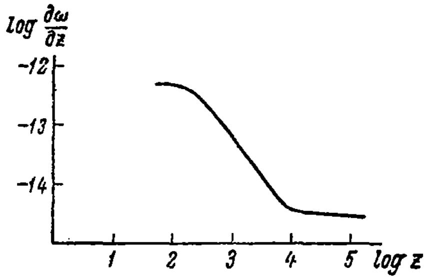


Figure 35. Dependence of differential solar rotation on depth

The above reasoning shows how important the role is of stratification for the dependence of the angular velocity of rotation of a star on astrographic latitude. It is obvious that stratification also is different in stars of different spectral classes and dimensions. Differential rotation, therefore, will also be expressed differently. However, here we cannot neglect the properties of the angle X, and we do not have such information for stars. A more detailed consideration of this problem, which falls in the field of stellar physics, is beyond the scope of this book.

If two stars have the same stratification in their subphotospheric layers, but the angles X for such stars are different, they will have a) different velocities of meridional circulation and b) the dependence of angular velocity of rotation on astrographic latitude will be greater for that star for which the angle X is greater. The first conclusion is derived directly from the circulation theorem, provided that the first term on the right-hand side is written in the form proposed by Silberstein. By limiting ourselves to this term alone, for the sake of simplicity and clarity, we can write the following expression for the circulation theorem for the case under consideration:

$$\oint v_\theta ds = \int_{t_0}^{t_1} \frac{1}{\rho^2} \frac{\partial p}{\partial n_1} \cdot \frac{\partial \rho}{\partial n_2} \sin \chi dt, \qquad (2.179)$$

hence, it must be concluded that there is a direct dependence of circulation (and this means velocity) on the angle X.

The derivation of the second conclusion in sufficiently general form is more complex. We did this in 1954, in the study already cited in connection with the problem of the velocity of change of latitude of the spot-forming zone on the sun (Ref. 80). The direct inspiration was an article by Gleissberg (Ref. 95) in which he made an attempt to

equalize the rates of movement of the spot-forming zone along meridians by use of the factor sin 2φ. If we denote by Δt the time expressed in solar rotations during which the spot-forming zone changes its heliographic latitude by 1°, it is found that the product $\Delta t \sin 2\varphi$ remains virtually unchanged. Such a result was obtained by Gleissberg using data for one 11-year cycle. We made a similar investigation using data for eight 11-year cycles. The results were more ambiguous than those obtained by Gleissberg, but it can be stated, nevertheless, that reduction by the use of the factor sin 2φ somewhat smooths the differences in $\dot{\varphi}$ for different φ. On the basis of his results Gleissberg assumed that

$$\dot{\varphi} = C \sin 2\varphi,$$

where C is a constant. On the other hand, if the Faye formula $\omega = a - b \sin^2\varphi$ is differentiated for latitude, we find

$$\frac{\partial\omega}{\partial\varphi} = -b \sin 2\varphi.$$

Hence Gleissberg draws the conclusion that $\dot{\varphi} \sim \frac{\partial\omega}{\partial z}$.

We attempted to substantiate this conclusion, proceeding on the basis of more general considerations. The Faye formula can be represented in the form of an approximate expression of the law of baroclinic rotation (Ref. 62);

$$\omega^2 = C_0\omega_0^2(r_0) + C_2\omega_2^2(r_0)P_2\mu, \tag{2.180}$$

where C_0 and C_2 are constants, ω_0 and ω_2 are dependent only on r (to be more precise, on a dimensionless value, being an argument of the Emden function). Therefore, in this particular case $\omega_0 = \omega_2$; then $\mu = \cos\theta$; $P_2(\mu) = \frac{3}{2}\mu^2 - \frac{1}{2}$ is the corresponding Legendre polynomial. Taking this into account, (2.180) will be

$$\omega^2 = \omega_0^2(r_0)\left(C_0 + \frac{3}{2}C_2\mu^2 - \frac{1}{2}C_2\right). \tag{2.181}$$

In the left-hand side of the Bjerknes-Rosseland formula we replace the essentially cylindrical coordinates l and z by the spherical coordinates r and θ. It possibly may be more correct to call them polar coordinates because the third coordinate, longitude or azimuth, is not involved due to rotational symmetry. The right-hand side of (2.147) is

used in expanded form as the product of the moduli of the gradients and the sine of the angle between the normals to the corresponding surfaces. Taking into account that $r = r_0$, after replacement of the variable and appropriate shortening we obtain

$$-2\omega \frac{\partial \omega}{\partial \theta} = F \sin \chi. \tag{2.182}$$

Here F is the product of the moduli of the gradients. From (2.181) we have

$$\omega = \omega_0 (r_0) \sqrt{C_0 + \frac{3}{2} C_2 \cos^2 \theta - \frac{1}{2} C_2}. \tag{2.183}$$

Substituting (2.182) into (2.181), we obtain

$$3\omega_0^2 (r_0) C_2 \sin 2\theta = F \sin \chi. \tag{2.184}$$

It therefore follows that with an increase of X, that is, with an increase of meridional circulation, with invariable θ and with other identical conditions (stratification) there is also an increase of $C_2\omega_0^2(r_0)$, that is, the term causing the dependence of the angular velocity of rotation on astrographic latitude. Thus, for a star, for which the effect of equatorial acceleration is more sharply expressed than for the sun, all other conditions (stratification, latitude) being equal, meridional circulation should occur more rapidly than on the sun (as a result of the large value X). The expression (2.184) includes the same factor sin 2θ "equalizing" meridional velocity. It is obvious that $\sin 2\theta = \sin 2\varphi$, since $\theta = \frac{\pi}{2} - \varphi$. The incomplete equalization of $\dot{\varphi}$ by the use of $\sin 2\varphi$ apparently can be attributed to certain changes of F within the 11-year cycle. Taking into account that during the course of the cycle the spot-forming zone moves from the high to the low latitudes, it is possible to postulate that the stratification of the subphotospheric layers in the high and low heliographic latitudes is somewhat different. In the next section we will describe one of the possible models giving the difference in the thickness of the convective zone as a function of heliographic latitude

Possibility of Considering the Convective Zone as a Boundary Layer

At the end of the section "Meridional Currents and Circulatory Instability", we discussed an investigation in which the convective zone was considered as a singular "stellar" boundary layer. Since we have turned to a consideration of this problem, it is necessary to recall first those initial theses which served as a basis for our work (Ref. 19) and then consider to what degree new data on this problem, especially the work of Biermann and Kippenhahn, mentioned above, have introduced substantial changes in our concepts with respect to these problems, i.e., whether it is possible subsequent to these studies to insist on the point of view which we expressed in the study mentioned (Ref. 19). Since this study (Ref. 19) was made using as a model the well-known study of N. Ye. Kochin for the earth's atmosphere, we first must consider to what degree it is proper to apply the same initial hypotheses to the sun.

In considering the earth's atmosphere (to be more exact, the troposphere as a boundary layer), Kochin proceeded on the assumption that the attachment condition is satisfied near the earth's surface, i.e., both the normal and tangential components of velocity should become equal to zero there. The satisfaction of such a condition causes no doubt in the case of the earth's atmosphere, since the earth's solid surface is present. In the case of the sun it is of course impossible to speak of a solid surface. However, the break in velocities at the base of the convective zone, that is, at its boundary with the inner zone of radiative equilibrium, should be so substantial that the latter can be considered virtually fixed in comparison with the first. This concept, which we expressed as early as 1958, has been confirmed completely by the latest investigations.

In actuality, according to Biermann (Ref. 94), meridional velocity in the convective zone is in order of magnitude 10^2 cm$\cdot$sec^{-1}. However, at the upper boundary of the inner zone of radiative equilibrium, according to Kippenhahn, it has the order of 10^{-6} or in the extreme case 10^{-5}, since the order of magnitude of vertical velocity will be $5 \cdot 10^{-7}$. Thus, in relation to the tangential component of velocity, the convective zone can be considered as a moving medium in relation to the fixed inner zone of radiative equilibrium. The matter should be still more striking with respect to vertical velocities, that is, the normal component of velocity. Whereas in the convective zone they have the order of 10^0 or 10^{-1}, in the zone of radiative equilibrium, as we have seen,

there is a vertical velocity of the order of $5 \cdot 10^{-7}$. Thus, for all practical purposes there is no penetration from the convective zone into the inner layers and back, with the exception of extremely slow diffusion or local nonstationary formations. After consideration of the above, the likening of the boundary of the convective zone and the inner zone of radiative equilibrium to a solid wall no longer appears so paradoxical.

In addition to satisfaction of the attachment conditions, the principal condition for the possibility of application of the boundary layer theory is a small thickness of the corresponding shell in comparison with its horizontal dimensions. For the earth's troposphere, the ratio of the vertical to the horizontal extent is in order of magnitude 0.003, provided the thickness of the troposphere is assumed to be the vertical distance within which the greater part of the air mass is concentrated. In the case of the sun, this relation, although somewhat less favorable, nevertheless fully permits the application of this theory. Actually, if the thickness of the convective zone is assumed to be even 0.2 of the solar radius, considering that its horizontal extent has the characteristic dimension of the order of six radii ($2\pi r_{\odot}$), we obtain for the ratio about 0.03; this does not take into account the distribution of density (that is, mass) in the convective zone. However, if we take into account what was done for the troposphere, the ratio of vertical to horizontal dimensions for the sun becomes still closer to that which applies for the earth.

In accordance with the characteristics assumed by N. Ye. Kochin for the earth's atmosphere, we will consider in sequence first stationary and zonal circulation, then zonal but not stationary circulation, and finally stationary but not zonal circulation. We recall that in stationary circulation the elements of motion are not dependent on time, and local (time) derivatives everywhere become equal to zero. Zonality of circulation, however, means that the elements of motion are not dependent on longitude. This means that although vertical and meridional motions can be present, they should be identical at all points on a particular parallel.

We will introduce the following system of spherical coordinates:

r—distance to the center of the celestial body $\theta = \frac{\pi}{2} - \varphi$, ψ is longitude. The projections of the velocity vector on the following directions: radius vector, tangent to the meridian and tangent to the parallel are denoted v_r, v_θ, v_ψ, respectively. The angular velocity of rotation of the sun, as usual, is denoted ω, and we will assume that it

is dependent neither on depth nor on latitude. Such an assumption is justified fully by the qualitative character of the reasoning; quantitatively it cannot influence the order of magnitude, which is most important from our point of view. As usual the acceleration of gravity is denoted g and the coefficient of dynamic viscosity is designated η. Under conditions of general circulation η is the exchange coefficient. The length of the radius of the lower base of the convective zone is denoted a and thickness of the zone is designated δ. To be more exact, by δ is meant the thickness of that part of the convective zone where the greater part of its mass is concentrated, i.e., the thickness of the lower part of the zone directly adjoining the inner zone of radiative equilibrium. This part, obviously, is at least several times smaller than the extent of the entire zone, so that it is correct to write the condition $\delta \ll a$. The general form of the continuity equation in Euler variables and in spherical coordinates will be

$$\frac{\partial\rho}{\partial t}+v_r\frac{\partial\rho}{\partial r}+\frac{1}{r}\frac{\partial\rho}{\partial\theta}+\frac{v_\theta}{r}\frac{\partial\rho}{\partial\theta}+\frac{v_\psi}{r\sin\theta}\cdot\frac{\partial\rho}{\partial\psi}+$$
$$+\rho\left(\frac{\partial v_r}{\partial r}+\frac{1}{r}\frac{\partial v_\theta}{\partial\theta}+\frac{1}{r\sin\theta}\frac{\partial v_\psi}{\partial\psi}+2\frac{v_r}{r}+\frac{\operatorname{ctg}\theta}{r}v_\theta\right)=0. \qquad (2.185)$$

In the case of zonal and stationary movement that we have considered, all of the terms containing derivatives of t and ψ disappear, and equation (2.185) is transformed into

$$\frac{\partial(\rho v_r r^2\sin\theta)}{\partial r}+\frac{\partial(\rho v_\theta r\sin\theta)}{\partial\theta}=0. \qquad (2.186)$$

Satisfaction of the attachment conditions means that

$$v_r=v_\theta=v_\psi=0 \text{ when } r=a. \qquad (2.187)$$

Integrating (2.186) in the limits from a to r, we have, taking (2.187) into account

$$v_r=-\frac{1}{\rho r^2\sin\theta}\int_a^r\frac{\partial(\rho v_\theta r\sin\theta)}{\partial\theta}dr. \qquad (2.188)$$

For now, in analysis of the continuity equation, we have assumed the conditions for stationary and zonal circulation, and also the attachment condition. We will now use the condition $\delta \ll a$. In considering general circulation, we should assume that a change in the elements of motion and the parameters of state occurs gradually. It is obvious that this cannot be said of those places where there is a discontinuity, that is, where there are fronts in the earth's atmosphere and active regions on the sun. In general, the changes should occur gradually, and the difference in values, such as meridional velocity Δv_θ

(here Δ denotes the difference and not the Laplacian), at two points can become commensurable with the value of this velocity when the distance between points is comparable with the radius of the celestial body.

In a study from the point of view which is now being presented, the continuity equation retains its validity, and as we will soon see, its investigation makes it possible to draw a series of important conclusions. However, if we considered the subphotospheric layers with discontinuities lying between them, it would be generally impossible to use the ordinary continuity equation; it would be necessary to replace it with a particular equation of conservation of mass for this case (Ref. 72).

Expressing the arc in abstract form, we find that the distance between two points for which the difference in the values of meridional velocity becomes commensurable with the velocity value itself will be equal to one radian. Thus, when $\Delta v_\theta = O(v_\theta)$ (the symbol O denotes the order of magnitude) and $\Delta\theta = O(1)$ we have

$$\frac{\partial v_\theta}{\partial \theta} = O(v_\theta) \tag{2.189}$$

(since the order of magnitude of the difference is equal to the order of magnitude of the derivative, with the order of magnitude of the difference of the values of the independent variable equal to unity). Similarly, for density we can write:

$$\frac{\partial \rho}{\partial \theta} = O(\rho). \tag{2.190}$$

Taking into account (2.189) and (2.190), and also that r differs from a by the value δ, we can obtain from expression (2.188) for the radial component of velocity

$$v_r = O\left(\frac{\delta u_s}{a}\right), \tag{2.191}$$

where v_s is the horizontal component of the velocity vector for either v_θ or v_ψ. It therefore follows that

$$O\left(\frac{v_r}{v_s}\right) = O\left(\frac{\delta}{a}\right). \tag{2.192}$$

This relation, applicable everywhere except in the immediate vicinity of the poles, where $\sin\theta$ in (2.188) is very small, is formulated as

follows: the order of magnitude of the ratio of the vertical to the horizontal component of velocity in movements of the scale of general circulation is the same as the order of magnitude of the ratio of the thickness of the circulation layer in which the greater part of the mass is concentrated to the radius of the corresponding celestial body. This relation has already been used in our discussions. Two comments must be made concerning its specific application to large-scale motions on the sun.

First, the displacement of the spot-forming zone shows that the difference in the value of meridional velocity becomes comparable to the mean value of this quantity when the distance between the corresponding points is less than the solar radius. The meridional velocity changes by a factor of 4 in a distance of 0.30 of the radius or by 7° per year for an arc corresponding to one radian. This is three times greater than the arc traveled annually by the spot-forming zone. However, since we are concerned with the order of magnitude, such a discrepancy is of no appreciable importance.

The second comment applies to vertical velocity. The vertical velocities of convection exceed by several orders of magnitude the horizontal velocity of general circulation. It would seem that this contradicts the relation established by formula (2.192) between the thickness of the zone in which the circulation occurs and the radius of the celestial body. In actuality, this is not true. The velocity of convection is characteristic for elements whose scale is appreciably less than the scale of general circulation, and even less than the scale of its horizontal elements.

From the continuity equation in integral form (2.188), it is possible to obtain the vertical component for movements of the scale of general circulation. However, ascending and descending elements associated with convection have a characteristic dimension not greater than $5 \cdot 10^{7}$ cm (this is the mean dimension of the mixing length (Ref. 52), that is, it should be even somewhat greater than the mean dimension of the convective elements themselves). However, the horizontal dimensions of the circulation elements, assuming approximately the same vertical extent, have a characteristic horizontal dimension of the order of 10^{10} cm (the extent of the active region as a certain dynamic entity (Refs. 8, 96)). The relation of the horizontal dimensions of the convective elements and elements of general circulation, therefore, has the same order of magnitude as the ratio of the mean transverse dimensions of a large cumulus cloud to the mean diameter of a low-pressure area.

It is completely clear that the presence of cumulus clouds in the earth's troposphere does not invalidate the applicability of the theory

of the planetary boundary layer to the earth's lower atmosphere. The large vertical velocities of ascending and descending convective elements on the sun can scarcely be of importance since these values apply to a relatively thin and rather high layer in the convective zone; near the base of the convective zone these velocities should be appreciably less, and they apparently do not significantly exceed the horizontal velocities of circulation. In brief, the same concepts can be discussed here that we already have touched upon in analyzing the recent investigation by Biermann.

We therefore conclude that the relation (2.192) is applicable under conditions of stationary and zonal circulation on the sun.

We will now consider zonal, nonstationary circulation. In this case the continuity equation will be

$$\frac{\partial(\rho r^2 \sin\theta)}{\partial t} + \frac{\partial(\rho v_r r^2 \sin\theta)}{\partial r} + \frac{\partial(\rho v_\theta r \sin\theta)}{\partial\theta} = 0. \qquad (2.193)$$

Hence

$$v_r = -\frac{1}{\rho r^2 \sin\theta}\int_a^r \left[\frac{\partial(\rho v_\theta r \sin\theta)}{\partial\theta} + \frac{\partial(\rho r^2 \sin\theta)}{\partial t}\right] dr. \qquad (2.194)$$

For comparison of the orders of magnitude of the horizontal and vertical velocities, we introduce the characteristic time interval τ, during which the elements of motion can change appreciably. It is apparent that the order of magnitude of the second term under the integral is related to the order of magnitude of the first as $\frac{a}{v_s \tau}$ (since $a \simeq r$). As indicated by N. Ye. Kochin (Ref. 79) for different forms of circulation in the earth's atmosphere, with the exception of those which have a very short lifetime, this ratio is of the order of several hundredths. As a result the second term under the integral can be discarded, and equation (2.194) is reduced to (2.192). Relation (2.192) is therefore also applicable in the earth's atmosphere for nonstationary circulation, since in this particular sense in the evaluation of the orders of magnitude it can be reduced to stationary circulation.

The situation will be different with respect to general circulation on the sun. In actuality, on the sun $a = 6\cdot 10^{10}$ cm, $v_s = 10^2$. As the characteristic time interval τ we use half a solar cycle, i.e., approximately five years. The greater the interval, the smaller will be the second term of the integrand (2.194) in comparison with the first. The

five years that we selected even constitute an excessively large value in that certain properties of general circulation change appreciably in as little as two years.

In the case of the numerical values used, $\frac{a}{v_a \tau} \approx 4$. Thus, even with such a large characteristic interval τ, the second term is 4 times greater than the first; if τ were assumed equal to two years, the second term would be 10 times greater than the first. It can therefore be concluded that relation (2.192) cannot be applied to general circulation on the sun if it is considered nonstationary.

Finally, we will consider stationary, but not zonal circulation. For this case the continuity equation will be

$$\frac{\partial (\rho v_r r^2 \sin\theta)}{\partial r} + \frac{\partial (\rho v_\theta r \sin\theta)}{\partial\theta} + \frac{\partial (\rho v_\psi r)}{\partial\psi}, \tag{2.195}$$

hence

$$v_r = -\frac{1}{\rho r^2 \sin\theta} \int_a^r \left[\frac{\partial (\rho v_\theta r \sin\theta)}{\partial\theta} + \frac{\partial (\rho v_\psi r)}{\partial\psi}\right] dr. \tag{2.196}$$

With θ not too close to zero, and taking into account that the orders of magnitude of the meridional and zonal components of the velocity vector are considered to coincide, we see that (2.196) is transformed into (2.188). Condition (2.192) is therefore satisfied for this case as well. When the difference in the orders of magnitude of v_θ and v_ψ is known, the order of magnitude of v_s obviously is determined by the larger of them.

A situation also can arise in which as a result of discarding of a value of a lesser order of magnitude we will have a formula of the type (2.188); but under the integral there will be a derivative of ψ. However, it is obvious that in such a case we can have reasoning similar to that above with respect to the relation of the orders of magnitude of v_s (in this case it will be v_ψ) and v_r. Therefore, condition (2.192) is satisfied for stationary general circulation of the sun, regardless of whether or not it is zonal. According to (2.192), the orders of magnitude of the vertical and horizontal velocities are related as the thickness of the layer in which the greater part of the circulating mass is related to the radius of the celestial body on which the circulation occurs.

We will now consider simplifications of the equations of motion which make it possible to obtain a second important relation concerning general circulation, that is, the relationship between the thickness of the layer, viscosity, Coriolis force and other forces. In contrast to conditions in the earth's atmosphere, we will also consider electromagnetic forces. In this study we will make no initial assumptions concerning the structure of the magnetic field, and only later will we introduce the hypothesis of a dipole field as the general magnetic field of the sun.

It is obvious that the introduction of electromagnetic forces into the equations of motion without a joint consideration of another highly important equation of magnetohydrodynamics—the induction equation—is palliative. However, the problem of the boundary layer in magnetohydrodynamics has not been developed sufficiently at the present time, and it would be premature here to attempt to apply incomplete theories to the problems of the boundary layer on rotating celestial bodies consisting of plasma. Since we have the far more modest objective of comparing the orders of magnitude of the forces entering into the equation of motion, we have felt it possible to confine ourselves to their consideration alone.

The equation of motion in vector form can be written:

$$\frac{\partial \bar{v}}{\partial t}+(\bar{v}\cdot\nabla)\bar{v}=\frac{1}{\rho}(\bar{i}\times\bar{H})-\frac{1}{\rho}\nabla p+\nu\nabla^2\bar{v}+\frac{\nu}{3}\nabla(\nabla\cdot\bar{v})+\frac{\bar{F}}{\rho}, \tag{2.197}$$

where $\bar{v}$ is velocity; t is time; $\bar{F}$ are external forces, other than viscosity and hydromagnetic forces; ρ is gas density; $\bar{i}$ is current density; $\bar{H}$ is magnetic field strength; p is gas pressure; and, ν is the coefficient of kinematic viscosity.

In accordance with the boundary layer theory as applicable to the particular case, it is necessary to convert to spherical coordinates. We will consider only stationary circulation, and we will first discuss a zonal case. The corresponding equations will be:

$$\begin{aligned}
&\rho\left(v_r\frac{\partial v_r}{\partial r}+\frac{v_\theta}{r}\frac{\partial v_r}{\partial\theta}-\frac{v_\theta^2+v_\psi^2}{r}\right)=-g\rho+2\omega\rho\sin\theta v_\psi-\frac{\partial p}{\partial r}+\\
&+\eta\left(\frac{\partial^2 v_r}{\partial r^2}+\frac{1}{r^2}\frac{\partial^2 v_r}{\partial\theta^2}+\frac{2}{r}\frac{\partial v_r}{\partial r}+\frac{\operatorname{ctg}\theta}{r^2}\frac{\partial v_r}{\partial\theta}-\frac{2}{r^2}\frac{\partial v_\theta}{\partial\theta}-\frac{2v_r}{r^2}-\right.\\
&\left.-\frac{2v_\theta\operatorname{ctg}\theta}{r^2}\right)+\frac{\partial\eta}{\partial r}\left(\frac{\partial v_r}{\partial r}-\frac{1}{r}\frac{\partial v_\theta}{\partial\theta}-\frac{2v_r}{r}-\frac{v_\theta\operatorname{ctg}\theta}{r}\right)+\\
&+\frac{1}{r}\frac{\partial\eta}{\partial\theta}\left(\frac{1}{r}\frac{\partial v_r}{\partial r}+\frac{\partial v_\theta}{\partial r}-\frac{v_\theta}{r}\right)+\frac{1}{3}\frac{\partial}{\partial r}\left[\eta\left(\frac{\partial v_r}{\partial r}+\frac{1}{r}\frac{\partial v_\theta}{\partial\theta}+\frac{2v_r}{r}+\right.\right.\\
&\left.\left.+\frac{v_\theta\operatorname{ctg}\theta}{r}\right)\right]+\sigma[H_\theta(v_\psi H_\psi+v_r H_r)-v_\theta(H_\psi^2+H_r^2)];
\end{aligned} \tag{2.198}$$

$$\rho\left(v_r\frac{\partial v_\theta}{\partial r}+\frac{v_\theta}{r}\frac{\partial v_\theta}{\partial\theta}+\frac{v_r v_\theta}{r}-\frac{v_\psi^2\operatorname{ctg}\theta}{r}\right)=-2\omega\rho\cos\theta v_\psi-\frac{1}{r}\frac{\partial p}{\partial\theta}+$$
$$+\eta\left(\frac{\partial^2 v_\theta}{\partial r^2}+\frac{1}{r^2}\frac{\partial^2 v_\theta}{\partial\theta^2}+\frac{2}{r^2}\frac{\partial v_\theta}{\partial r}+\frac{\operatorname{ctg}\theta}{r^2}\frac{\partial v_\theta}{\partial\theta}-\frac{v_\theta}{r^2\sin^2\theta}\right)+$$
$$+\frac{\partial\eta}{\partial r}\left(\frac{1}{r}\frac{\partial v_r}{\partial\theta}+\frac{\partial v_\theta}{\partial r}-\frac{v_\theta}{r}\right)+\frac{1}{r}\frac{\partial\eta}{\partial\theta}\left(\frac{1}{r}\frac{\partial v_\theta}{\partial\theta}-\frac{\partial v_r}{\partial r}-\frac{v_\theta\operatorname{ctg}\theta}{r}\right)+$$
$$+\frac{1}{3r}\frac{\partial}{\partial\theta}\left[\eta\left(\frac{\partial v_r}{\partial r}+\frac{1}{r}\frac{\partial v_\theta}{\partial\theta}+\frac{2v_r}{r}+\frac{v_\theta\operatorname{ctg}\theta}{r}\right)\right]+$$
$$+\sigma\left[H_\theta(v_\psi H_\psi+v_r H_r)-v_\theta(H_\psi^2+H_r^2)\right]; \qquad (2.199)$$

$$\rho\left(v_r\frac{\partial v_\psi}{\partial r}+\frac{v_\theta}{r}\frac{\partial v_\psi}{\partial\theta}+\frac{v_r v_\psi}{r}+\frac{v_\theta v_\psi\operatorname{ctg}\theta}{r}\right)=-2\omega\rho\cos\theta v_\theta+$$
$$+\eta\left(\frac{\partial^2 v_\psi}{\partial r^2}+\frac{1}{r^2}\frac{\partial^2 v_\psi}{\partial\theta^2}+\frac{2}{r}\frac{\partial v_\psi}{\partial r}+\frac{\operatorname{ctg}\theta}{r^2}\frac{\partial v_\psi}{\partial\theta}-\frac{v_\psi}{r^2\sin^2\theta}\right)+$$
$$+\frac{\partial\eta}{\partial r}\left(\frac{\partial v_\psi}{\partial r}-\frac{v_\psi}{r}\right)+\frac{1}{r}\frac{\partial\eta}{\partial\theta}\left(\frac{1}{r}\frac{\partial v_\psi}{\partial\theta}-\frac{v_\psi}{r}\operatorname{ctg}\theta\right)+$$
$$+\sigma\left[H_\psi(v_r H_r+v_\theta H_\theta)-v_\psi(H_r^2+H_\theta^2)\right], \qquad (2.200)$$

where H_r, H_θ and H_ψ are the magnetic field components along the r-, θ- and ψ-axes, respectively, and σ is conductivity.

By analogy with (2.192) it is possible to write $\frac{\partial^2 v_r}{\partial\theta^2}=O(v_r)$, since $\frac{\partial^2 v_s}{\partial\theta^2}=\frac{\partial}{\partial\theta}\frac{\partial v_s}{\partial\theta}$, but $\frac{\partial v_s}{\partial\theta}=O(v_s)$, and exactly the same, $\frac{\partial v_r}{\partial\theta}=O(v_r)$, $\frac{\partial^2 v_r}{\partial\theta^2}=O(v_r)$. Since $v_r=O\left(\frac{\delta v_s}{a}\right)$, then $\frac{\partial v_r}{\partial\theta}=O\left(\frac{\delta}{a}v_s\right)$, $\frac{\partial^2 v_r}{\partial\theta^2}=O\left(\frac{\delta}{a}v_s\right)$. The derivatives for the direction of the vertical will be appreciably different. Here the changes v_s or v_r, equal in order of magnitude to the value itself, should occur after a distance of the order of δ has been covered, since in accordance with the attachment conditions the velocities are equal to zero when r = a, and attain their normal values at the level r = a + δ. The attachment conditions are boundary conditions at the lower boundary. At the upper boundary we have a distribution of velocities associated with the motions of sunspot groups.

Consequently,

$$\frac{\partial v_s}{\partial r}=O\left(\frac{v_s}{\delta}\right),\quad \frac{\partial^2 v_s}{\partial r^2}=O\left(\frac{v_s}{\delta^2}\right),\quad \frac{\partial v_r}{\partial r}=O\left(\frac{v_r}{\delta}\right),\quad \frac{\partial^2 v_r}{\partial r^2}=O\left(\frac{v_r}{\delta^2}\right). \qquad (2.201)$$

These expressions, with (2.192) taken into account, make it possible to discard terms in equations (2.197)-(2.199) containing the factors $\frac{\delta}{a}$ and $\frac{\delta}{a^2}$. Introducing the notation

$$\bar{p}=p-\frac{1}{3}\eta\left(\frac{\partial v_r}{\partial r}+\frac{1}{r}\frac{\partial v_\theta}{\partial \theta}+\frac{v_\theta \operatorname{ctg}\theta}{r}\right),$$

the equations can be reduced to the form:

$$\rho\left(v_r\frac{\partial v_\theta}{\partial r}+\frac{v_\theta}{r}\frac{\partial v_\theta}{\partial \theta}-\frac{v_\psi^2 \operatorname{ctg}\theta}{r}\right)=-2\omega\rho\cos\theta v_\psi-\frac{1}{r}\frac{\partial \bar{p}}{\partial \theta}+$$
$$+\eta\frac{\partial^2 v_\theta}{\partial r^2}+\frac{\partial \eta}{\partial r}\frac{\partial v_\theta}{\partial r}+\sigma\left[H_\theta\left(v_\psi H_\psi+v_r H_r\right)-v_\theta\left(H_\psi^2+H_r^2\right)\right], \tag{2.202}$$

$$\rho\left(v_r\frac{\partial v_\psi}{\partial r}+\frac{v_\theta}{r}\frac{\partial v_\psi}{\partial \theta}+\frac{v_\psi v_\theta \operatorname{ctg}\theta}{r}\right)=-2\omega\rho\cos\theta v_\theta+$$
$$+\eta\frac{\partial^2 v_\psi}{\partial r^2}+\frac{\partial \eta}{\partial r}\frac{\partial v_\psi}{\partial r}+\sigma\left[H_\psi\left(v_r H_r+v_\theta H_\theta\right)-v_\psi\left(H_r^2+H_\theta^2\right)\right]. \tag{2.203}$$

In these equations it is possible to compare the orders of magnitude of inertia and Coriolis force. The order of magnitude of the first is $\frac{\rho v_s^2}{a}$, and of the second,— $\omega\rho v_s$. The ratio is $\frac{\rho v_s^2}{a}:\omega\rho v_s=\frac{v_s}{a\omega}$. For the earth this ratio is 0.03. On the sun we have still more favorable conditions for neglecting inertia in comparison with Coriolis force. This should be true on the basis of similarity theory, since the scale of general circulation on the sun is greater than on the earth. Assuming as before that $v = 10^2$, $a = 6\cdot 10^{10}$, and $\omega = 3\cdot 10^{-6}$, we have $\frac{v_r}{a\omega} = 0.0005$. Thus, in considering motions of the scale of the general circulation of the sun it is possible to neglect inertial forces completely. We note in passing that the result obtained above during investigation of the continuity equation as it concerns the satisfaction of condition (2.192) only for stationary circulation, also is related to more considerable scales of movement on the sun, and at the same time is related to a lesser velocity of general circulation than characteristic of terrestrial conditions.

Neglecting inertial forces, we can write:

$$\eta\frac{\partial v_\theta}{\partial r^2}+\frac{\partial \eta}{\partial r}\frac{\partial v_\theta}{\partial r}+2\omega\rho\cos\theta v_\psi=\frac{1}{r}\frac{\partial \bar{p}}{\partial \theta}-\sigma\left[H_\theta\left(v_\psi H_\psi+\right.\right.$$
$$\left.\left.+v_r H_r\right)-v_\theta\left(H_\psi^2+H_r^2\right)\right], \tag{2.204}$$

$$\eta \frac{\partial^2 v_\psi}{\partial r^2} + \frac{\partial \eta}{\partial r}\frac{\partial v_\psi}{\partial r} - 2\omega\rho \cos\theta v_\theta + \sigma [H_\psi (v_r H_r + \\ + v_\theta H_\theta) - v_\psi (H_r^2 + H_\theta^2) = 0. \tag{2.205}$$

If the electromagnetic forces in the second of these equations are appreciably less than the others, (2.205) assumes the form

$$\eta \frac{\partial^2 v_\psi}{\partial r^2} + \frac{\partial \eta}{\partial r} \cdot \frac{\partial v_\psi}{\partial r} - 2\omega\rho \cos\theta v_\theta = 0. \tag{2.206}$$

In order of magnitude this will be

$$\eta \frac{v_s}{\delta^2} - \omega\rho \cos\theta v_s = 0, \tag{2.207}$$

hence

$$\delta = \sqrt{\frac{\eta}{\omega\rho \cos\theta}}. \tag{2.208}$$

This formula is known from the theory of general circulation of the atmosphere (Ref. 79). If viscous forces in equation (2.206) are appreciably less than the other forces, it will be impossible to determine the order of magnitude of the thickness of the layer (zone). However, under all other circumstances it is possible to obtain from (2.204) the order of magnitude of the meridional pressure gradient

$$\frac{1}{r} \cdot \frac{\partial \bar{p}}{\partial \theta} = \eta \frac{v_s}{\delta^2} + \omega\rho \cos\theta v_s + \sigma [H_\theta (v_s H_\psi + v_r H_r) - \\ - v_s (H_\psi^2 + H_r^2)]. \tag{2.209}$$

The use of $\bar{p}$ instead of p obviously cannot exert an influence on the order of magnitude, since in the expression

$$\bar{p} = p - \frac{1}{3}\eta\left(\frac{\partial v_r}{\partial r} + \frac{1}{r}\frac{\partial v_\theta}{\partial \theta} + \frac{v_\theta \operatorname{ctg}\theta}{r}\right)$$

the second term has an order of magnitude $\frac{\nu\rho v_s}{r}$ not greater than 10^3, and the first 10^{12}. Until now we have imposed no limitations on the structure of the magnetic field determining the electromagnetic forces. However, it is possible to assume that the radial component H_r has the same order of magnitude as the other components. This assumption will hold true for a broad class of fields; it is, for example, correct for a dipole field, except for points close to the pole. If this assumption is made, and it is taken into account that $v_r \ll v_s$, the expression for

electromagnetic forces is simplified. Formulas (2.198-2.200) assume the following form:

$$\eta\frac{\partial^2 v_\theta}{\partial r^2}+\frac{\partial\eta}{\partial r}\frac{\partial v_\theta}{\partial r}+2\omega\rho\cos\theta v_\psi=\frac{1}{r}\frac{\partial\bar{p}}{\partial\theta}-$$
$$-\sigma[H_\theta H_\psi v_\psi-v_\theta(H_\psi^2+H_r^2)],\tag{2.210}$$

$$\eta\frac{\partial^2 v_\psi}{\partial r^2}+\frac{\partial\eta}{\partial r}\frac{\partial v_\psi}{\partial r}-2\omega\rho\cos\theta v_\theta+$$
$$+\sigma[H_\psi H_\theta v_\theta-v_\psi(H_r^2+H_\theta^2)]=0,\tag{2.211}$$

$$\frac{1}{r}\frac{\partial\bar{p}}{\partial\theta}=\eta\frac{v_s}{\delta^2}+\omega\rho\cos\theta v_s+$$
$$+\sigma v_s[H_\theta H_\psi-(H_\psi^2+H_r^2)].\tag{2.212}$$

If none of the forces can be neglected in equation (2.205), the thickness of the zone ("stellar" boundary layer) will be

$$\delta=\sqrt{\frac{\eta}{\omega\rho\cos\theta-\sigma[H_\psi H_\theta-(H_r^2+H_\theta^2)]}}.\tag{2.213}$$

Formula (2.213) gives the order of magnitude of the thickness of the layer, assuming the admissibility of the simplification which was made with respect to the components of the magnetic field and velocities. Formula (2.212) gives the order of magnitude of the meridional pressure gradient under these same assumptions; finally, formula (2.209) gives the order of magnitude of the pressure gradient in a general case.

We will now consider the case of azonal circulation. We will not write equations which are still more unwieldy than (2.198)-(2.220). Again using a small value of the quantities $\frac{\delta}{a}$ and $\frac{v_s}{a\omega}$, it is possible to reduce the more complex equations to the form

$$\left.\begin{aligned}&2\omega\rho\cos\theta v_\psi-\frac{1}{r}\frac{\partial\bar{p}}{\partial\theta}+\frac{\partial}{\partial r}\left(\eta\frac{\partial v_\theta}{\partial r}\right)+\sigma[H_\theta H_\psi v_\psi-\\&\qquad-v_\theta(H_\psi^2+H_r^2)]=0,\\&-2\omega\rho\cos\theta v_\theta-\frac{1}{r\sin\theta}\frac{\partial\bar{p}}{\partial\psi}+\frac{\partial}{\partial r}\left(\eta\frac{\partial v_\psi}{\partial r}\right)+\\&\qquad+\sigma[H_\psi H_\theta v_\theta-v_\psi(H_r^2+H_\theta^2)]=0.\end{aligned}\right\}\tag{2.214}$$

The first of these equations does not differ from (2.210); from the second it is possible to obtain the order of magnitude of the zonal pressure gradient. In this case, there are no equations similar to (2.211), since the zero in the right-hand side of equation (2.211) is in fact related to the zonality condition. It is easy to see that the same assumptions have been made in equation (2.214) with respect to electromagnetic forces as were made in (2.210). If the magnetic field is considered a dipole field, and it is assumed for simplicity that its axis coincides with the sun's axis of rotation, the field components along the axes of the spherical coordinates will be:

$$H_r = H_p \cos\theta, \tag{2.215}$$

$$H_\theta = \frac{H_p}{2}\sin\theta, \tag{2.216}$$

$$H_\psi = 0. \tag{2.217}$$

By substituting these expressions into formulas (2.210) and (2.211) we obtain

$$\eta\frac{\partial^2 v_\theta}{\partial r^2} + \frac{\partial\eta}{\partial r}\frac{\partial v_\theta}{\partial r} + 2\omega\rho\cos\theta v_\psi = \frac{1}{r}\frac{\partial\bar{p}}{\partial\theta} + \sigma(v_\theta H p^2\cos^2\theta), \tag{2.218}$$

$$\eta\frac{\partial^2 v_\psi}{\partial r^2} + \frac{\partial\eta}{\partial r}\frac{\partial v_\psi}{\partial r} - 2\omega\rho\cos\theta v_\theta -$$
$$-\sigma\left[v_\psi\left(H_p^2\cos^2\theta + \frac{H_p^2}{4}\sin^2\theta\right)\right] = 0. \tag{2.219}$$

It follows from (2.219) that

$$\delta = \sqrt{\frac{\eta}{\omega\rho\cos\theta \pm \sigma\left(H_p^2\cos^2\theta + \frac{H_p^2}{4}\sin^2\theta\right)}}. \tag{2.220}$$

The meridional gradient is determined from (2.218). In the case of azonal circulation it is possible to determine the zonal gradient from a formula that follows from (2.214), after substituting the relations (2.215)-(2.217) into it. This will obviously be

$$\frac{1}{r\sin\theta}\frac{\partial\bar{p}}{\partial\psi} = -2\omega\rho\cos\theta v_\theta + \frac{\partial}{\partial r}\left(\eta\frac{\partial v_\psi}{\partial r}\right) -$$
$$-\sigma v_\psi\left(H_p\cos^2\theta + \frac{H_p^2}{4}\sin^2\theta\right). \tag{2.221}$$

We will now proceed to numerical evaluations. It is most objective to obtain such an evaluation for a purely hydrodynamic case. We

replace $\frac{\eta}{\rho}$ in formula (2.208) by the coefficient of kinematic turbulent viscosity K, and take into account that $O(K) = O(vL)$, where v is the velocity of circulation, and L is its characteristic horizontal dimension. Since $v = 10^2$ cm·sec^{-1}, $L = 10^{11}$ cm, $K = 10^{13}$ cm^2·sec^{-1}. Taking into account that the order of magnitude of ω also is maintained at the level of the base of the convective zone, we have $\delta = 3 \cdot 10^9$ cm, which agrees well with the thickness of the convective zone obtained by completely different methods (on the basis of thermodynamic considerations, sunspot depth, etc.).

This conclusion is the first a posteriori confirmation of the correctness of the proposed model. We note for refinement that the investigation was made for $\theta = 60^\circ$, that is, $\cos\theta = 1/2$. Since the denominator of the radicand in (2.220) in principle contains 2, as a result we have there simply ω. In the sense of order of magnitude what has been said is of course without significance.

We will now estimate the value of the meridional and zonal pressure gradients. We will again proceed on the basis of purely hydrodynamic considerations. Then we obtain the following from (2.208), with δ assumed equal to $5 \cdot 10^9$ cm for sake of clarity: the meridional gradient will be $1.4 \cdot 10^{-6}$. In this case the term taking viscosity into account is of the order of 10^{-7} and the Coriolis force is 10^{-6}. It can therefore be concluded that the "wind" is almost geostrophic (or to be more exact, "heliostrophic"). This explains why the results of the geostrophic approximation used in Ref. 80 are satisfactory.

The value of the zonal gradient at a particular polar distance will be very close to the meridional value. It will be 10^{-6}. If it is assumed (and this is confirmed to a certain degree by observations) that the zonal velocities are somewhat higher than the meridional velocities, as a result of its geostrophic character of movement the meridional gradient will be greater than the zonal gradient, a hypothesis which we used in Ref. 80.

We will now consider the corresponding equations when electromagnetic forces are taken into account. Under conditions of a dipole field the thickness of the layer (zone) is given by formula (2.220). For computation of the order of magnitude of the electromagnetic forces entering into this expression we take into account that according to

Table 17 the gas kinetic conductivity σ at a depth of $5 \cdot 10^9$ cm is equal to $5 \cdot 10^{-6}$ CGSM units. The dipole field is considered to increase in depth in accordance with the inverse cubes law. This means that when the field strength at the surface is 1 oe, at a depth of $5 \cdot 10^9$ cm it will be approximately 1.2 oe.

Making the computations, we obtain for the term taking electromagnetic forces into account in (2.220) a value of $1.7\ 10^{-6}$, at the same time obtaining a value of $1.2 \cdot 10^{-8}$ for the term taking Coriolis force into account in this same formula (in the latter case the factor v_s is absent). We find, therefore, that the term taking Coriolis force into account can be discarded. However, physically such rejection is without sense, since in this case it is found that the thickness of the layer is determined only through the forces of resistance.

We can therefore repeat the conclusion that we drew in Ref. 19: under conditions of gas kinetic conductivity, it is impossible to have zonal circulation in the convective zone, considered as a boundary layer. However, from the equation for azonal circulation, it is impossible to find the meridional and zonal gradients for this case, since the thickness of the zone is unknown (the two corresponding equations contain three unknowns: thickness of the zone, the meridional gradient, and the zonal gradient). If we take the first of these unknowns, proceeding on the basis of the theory of an extended convective zone, for example, we find that such a thickness of the zone cannot be correlated with the value of gasdynamic conductivity. This situation becomes clear if we take into account the coincidence of the orders of magnitude of the zonal and meridional gradients, on the one hand, and what we have said above concerning the contradiction between the value of the electromagnetic forces and Coriolis force for a zonal case on the other.

Thus, in an azonal circulation the problem of determining the thickness of the layer and the gradients becomes ambiguous. Assuming zonal circulation, however, we obtain the correct order of magnitude of both the thickness of the convective zone and the meridional pressure gradient. The pole-to-equator temperature drop, computed on the basis of the meridional gradient determined from the equation for zonal circulation, is equal to about 0.2°, which agrees very well with 0.3°, which is obtained from a consideration of the dependence of temperature on the acceleration of gravity, which changes with latitude. There is basis for assuming that the general circulation on the sun is zonal.

In actuality, the displacement of the spot-forming zone from the high to the low latitudes during the 11-year cycle occurs uniformly at all longitudes. In the case of virtually "heliostrophic" movement this means a constancy of the zonal gradient at different longitudes, i.e., the condition of zonality of circulation. Thus, in a purely hydrodynamic case everything is more or less favorable. The picture changes sharply when an allowance is made for electromagnetic forces, whose order of magnitude appreciably exceeds the order of magnitude of the Coriolis force. In this case the determination of the thickness of the convective zone, considered as a boundary layer, becomes impossible, and it is also impossible to determine the meridional gradient.

Conclusions

We will now attempt to summarize certain results of this chapter.

1. At present, most investigators are inclined to assume that the convective zone is extensive and has a thickness of 0.1-0.2 of the solar radius.

2. In the upper part of the convective zone stratification is very unstable, and there is a well-expressed convection. The difference between the real and adiabatic gradients is relatively great. In the deeper layers the stratification becomes almost adiabatic.

3. The vertical velocity of convection increases rapidly with depth, and then slowly decreases. The transport of energy in the entire extent of the zone is accomplished for the most part by convection.

4. At the lower boundary of the convective zone, the transition to conditions characteristic of radiative equilibrium occurs rather suddenly.

5. The entire convective zone is characterized by high viscosity associated with turbulent phenomena. It is impossible to agree fully with the Biermann concept that the coefficient of turbulent viscosity is separated into isotropic and monotropic components. It must always be remembered which turbulent elements are responsible for exchange.

6. The conclusion that there is a relatively slow increase of the angular velocity of solar rotation with depth in the convective zone is confirmed by an estimate of the balance of angular momentum from an equation of the Jeans type, except for the convective zone. The most natural mechanism of differential rotation is the nature of a baroclinic fluid in combination with an excess of temperature and pressure in the high heliographic latitudes in comparison with the equatorial latitudes. It is scarcely satisfactory to attribute the higher temperature in the high latitudes to flattening, since in this case differential

rotation should occur down to the deepest layers of the sun. Moreover, the virtual invariability of the position of the active longitudes indicates otherwise. It is not impossible that the reason for the certain temperature asymmetry is magnetohydrodynamic waves that transport part of the energy and propagate differently along the axis of the general magnetic field and in a transverse direction.

7. In the case of zones of radiative equilibrium the dependence of the change of angular velocity with distance to the equatorial plane on the linear velocity of rotation and on conditions of baroclinic stratification is given by the well-known Bjerknes-Rosseland theorem. In the convective zone, the equality given by the above-mentioned theorem is transformed according to Biermann into an inequality. The change of angular velocity with latitude naturally should occur there more slowly than in regions where turbulent viscosity is not effective. However, the effect of the latter is not significant.

8. According to Biermann, the velocity of meridional circulation is of the order of 10^2 cm$\cdot$sec^{-1}. Approximately the same order of velocity is obtained for the photosphere (somewhat less), provided we use simple equations of geostrophic motion. The basis for such a possibility (use of the mentioned equations) is elementary computations of the forces of the pressure gradient, based on the temperature drop (and the pressure drop following it, or vice versa) between the high and low solar latitudes, and the deflecting force associated with solar rotation.

9. According to Kippenhahn, the velocity of meridional circulation increases toward the outer layers in proportion to the density decrease. At the level of the upper layers of the inner zone of radiative equilibrium this velocity, when computations are made by the Kippenhahn method, differs little from the velocity of Eddington currents. The use of the equations of geostrophic motion for this level and deeper layers apparently loses sense, since these layers already lie lower than those in which there is a differentiation of circulation, and there is a zonal gradient causing meridional movements at shallower levels.

10. The existence of active longitudes is an argument in favor of differentiation of circulation and a zonal pressure gradient giving rise to meridional circulation. At heliographic latitudes greater than 15° the role of the zonal gradient is especially great; under certain conditions in the lower latitudes (if the equator-pole temperature drop is greater than follows from the theory of radiative equilibrium, and approaches the value found by Minnaert) it is possible to have movement along a meridional gradient. The displacement of formations belonging to higher layers of the sun's atmosphere (filaments) into higher latitudes in the final stages of the lifetime of active regions can be

related to some degree also to circulatory phenomena. This displacement can characterize the upper opposite branch of a large-scale circulatory center, resulting from the longitudinal differentiation of circulation. If this is true, circulatory centers on the sun should be oriented at a considerable angle to the sun's surface.

11. The considerable velocity of circulation in the convective zone, according to Biermann, and its small value in the inner zone of radiative equilibrium, according to Kippenhahn, suggests the presence of a velocity discontinuity near the lower boundary of the convective zone. Therefore, to the two forms of instability characteristic of the subphotospheric layers of the sun, gravitational and rotational, we can add a third form—circulatory instability. As an intermediate model between the models of a smooth change of velocity and a discontinuity, it is possible to consider a model of a boundary layer ("stellar") occupying the entire sun. In this case, it is assumed initially that the attachment conditions are satisfied at the boundary of the zone of radiative equilibrium and the zone of convection. The results obtained for conditions of stationary zonal circulation have made it possible to explain the order of magnitude of such values as the thickness of the convective zone (prior to the hydrodynamic approach, there actually was no way to substantiate its extent and the latter was established on the basis of purely thermodynamic considerations and certain concepts from mixing length theory) and the value of the meridional pressure gradient. In this case, however, it must be assumed that electromagnetic forces are not operative.

12. The degree of dependence of the angular velocity of solar rotation on heliographic latitude is determined by two factors—the angle of inclination of isobaric surfaces to isosteric surfaces and the moduli of the pressure and specific volume gradients. The first characteristic is in all probability almost constant for a particular star, while the second is strongly dependent on depth. Already in the lower photosphere the value $\frac{\delta\omega}{\delta z}$, determined from the Bjerknes-Rosseland theorem, in the case of a constant angle of inclination X of isosurfaces, is an order of magnitude less than that in the higher layers of the photosphere.

In the inner zone of radiative equilibrium, the product of the moduli of the gradients of these values is four orders of magnitude less than in the photosphere. Because of this, and taking into account conclusion No. 5, it can be said that solar rotation in the convective zone and in the inner zone of radiative equilibrium is virtually barotropic.

13. In the case of two stars having the same stratification, that star on which the angle X is the larger will have meridional circulation

of a greater velocity. The ascending branch of activity cycles in such a star will be shorter. By analogy with the sun, this should mean that the cycles will be more clearly expressed. It is not impossible that in stars with a considerable angle X the activity cycles can lead to variations in luminosity, i.e., stellar variability.

CHAPTER 3
ENERGY SOURCES OF SOLAR ACTIVITY

Section 1. General Comments

The answer to the problem of the source of the unquestionably enormous energy that is expended in solar activity processes, and in the shortwave, corpuscular, radio, and other radiation that exceeds the level of the quiet sun, and which has been classified by Biermann as nonthermal (Ref. 1), at first glance appears to be quite elementary: the source of solar activity is all that same nuclear energy that is responsible for the activity of the quiet sun. However, beginning with the time of Wolf, and to the present time, attempts have continued to demonstrate what might be called the exogenic nature of this energy. We have discussed this problem briefly in Chapter 2. It is still more important that, while admitting the endogenic character of the energy of solar activity, we are still quite far from a solution to the problem of the specific forms in which the sun's internal energy is transformed into the different forms of energy characteristic of its activity.

In this connection we feel it rational to consider first the interrelationship between the different manifestations of solar activity energy, including here the energy of general circulation on the sun. It is then desirable to proceed to an estimate of the energy involved in exogenic factors. In concluding this chapter an attempt will be made to estimate as a whole the energy of active solar radiation. This should provide the basis for the three chapters that follow, and are devoted to manifestations of solar activity.

Section 2. Comparison of the Magnetic Energy of Sunspots, the Energy of Differential Solar Rotation and its General Circulation

In the preceding chapters we have indicated repeatedly the relationship existing between the laws of change of solar activity in the 11-year cycle, changes in the latitude of spot formation in this cycle, and solar rotation. We now note some of the basic conclusions: (a) the more intense the 11-year cycle, the higher is the heliographic latitude at which the first spot groups belonging to this cycle appear (Ref. 2); (b) the more intense the cycle, the higher is the heliographic latitude at which its maximum occurs (Ref. 3) (this conclusion for the most part is associated with the fact that a more intense cycle begins at higher

heliographic latitudes and has a shorter ascending branch (Ref. 4)); (c) in the more intense cycles the rate of creeping of the spot-forming zone from the high to the low latitudes is somewhat greater (Ref. 4); (d) since 11-year cycles end at virtually the same latitude, regardless of the height of their maxima, in the case of the more intense cycles, the spot-forming zone passes along a great circle (Ref. 4); and, (e) at relatively low heliographic latitudes a specific Wolf number corresponds to a particular latitude, regardless of the intensity of the 11-year cycle (Ref. 2).

With respect to the law of solar rotation, in this case there is a relationship between the "expressiveness" of differential rotation and the rate of creeping of the spot-forming zone (Ref. 5), and therefore with the intensity of the cycle. Unfortunately, for now there is no directly observed confirmation of the latter property (a relationship with intensity).

In his studies, Biermann pointed out the role of factors which cause differential rotation in the initial formation of the sun's general magnetic field (Ref. 6). This suggested that differential solar rotation can serve as the source of energy which is transformed in spots (or under them) into magnetic energy. Alfvén (Ref. 7) has spoken out against such a hypothesis; we will review his reasoning briefly. Proceeding on the basis that bipolar groups are the principal structural units possessing a strong magnetic field, Alfvén writes the following expression for magnetic energy:

$$W_s = \frac{H^2}{8\pi}\tau. \tag{3.1}$$

Here $\frac{H^2}{8\pi}$ is the density of magnetic energy, τ is the volume occupied by a bipolar group. Assuming that the lines of force form a subphotospheric tube connecting both spots of the group, and length of the tube is d, this volume can be described as

$$\tau = A_s d, \tag{3.2}$$

where A_s is spot area. Alfvén neglects the volume of the central part (this simplification should not affect the order of magnitude). Assuming the mean strength of the magnetic field of a spot to be 3,000 gauss and the mean spot area of the sun to be $3 \cdot 10^{19}$ cm^2, which corresponds to a total spot area of 1,000 millionths of a hemisphere, Alfvén obtains as the magnetic energy of spots

$$W = \frac{(3000)^2 \cdot 10^{10} \cdot 3 \cdot 10^{19}}{8\pi} \approx 10^{35}$$

Alfvén assumes the mean lifetime of a group to be 0.1 year and believes that the magnetic energy also is dissipated during this time. The magnetic energy dissipating during a year therefore will be

$$P = 10^{36} \text{ ergs/yr.}$$

According to Alfvén the kinetic energy of differential solar rotation is 10^{39}-10^{40} ergs. After comparing the value P and the energy of differential rotation, Alfven concludes that the latter would suffice to supply magnetic energy to sunspots for only 5,000 years. He feels that it is improbable that the energy of differential rotation could be supplemented in such a short period. Alfven therefore regards as untenable the hypothesis that there is a relationship between differential solar rotation and the magnetic energy of sunspots. However, an analysis of the results obtained by Alfvén shows that the problem is more complex (Ref. 8). Computations show that an energy of differential rotation of 10^{39}-10^{40} ergs will have this value in a case in which the polar retardation observed in the photosphere involves a layer with a thickness of about 200,000 km.

As pointed out in the preceding chapter, polar retardation decreases rapidly with depth, and more or less appreciable differential rotation is therefore a property of layers very close to the surface. Taking this factor into account, the total energy of differential rotation in a zone with a thickness of about 200,000 km under the photosphere is $8 \cdot 10^{35}$ ergs, which is even less than the estimate of magnetic energy made by Alfvén. At the same time, it must be remembered that the determined value of the energy of differential rotation is even exaggerated somewhat, since for the zone of convection the maximum of the possible values of "equatorial acceleration" has been used. Thus, if the assumption made in the preceding chapter that there is a constancy of the angle of inclination of isobaric surfaces to isosteric surfaces is correct, that is, if "differential rotation" actually decreases with depth approximately as indicated in Figure 35, the energy of differential rotation in general is insufficient to supply the magnetic energy of sunspots.

By separate study of the energy of differential rotation of the inner zone of radiative equilibrium, the convective zone, and the photosphere, it can be seen that, if for the first of these layers only its peripheral part, with a thickness of about 90,000 km, is limited, for each of the three mentioned layers, the energy of differential rotation will be: for the outer part of the inner zone of radiative

equilibrium—$6.0 \cdot 10^{32}$ ergs; for the convective zone—$8.2 \cdot 10^{35}$ ergs; and, for the photosphere—$7.9 \cdot 10^{33}$ ergs. As a sum we obtain $8.3 \cdot 10^{35}$ ergs, which is less than the total energy of the magnetic fields of spots estimated by Alfvén.

It, therefore, cannot be said that the sole source of the magnetic energy of solar activity is differential solar rotation. We will now compare the energy of differential rotation with the magnetic energy involved in fluctuations of solar activity, that is, in the excesses above the mean background for a particular phase of the solar cycle. Computation of the magnetic energy in fluctuations was done by R. N. Ikhsanov (Ref. 9). Table 24, taken from his published studies, but with corrections kindly communicated by the author, and in somewhat abbreviated form shows the epochs of the maximum of certain fluctuations and the magnetic energy of the spots which they caused.

It can therefore be seen that the mean energy of a fluctuation is an order of magnitude less than the energy of differential rotation. Numerically, the second, therefore, can be the source of the first. However, even disregarding the fact that it is strange why the energy of the "background" comes from one source and the energy of fluctuations from another, the hypothesis that differential rotation is the source of sunspot energy can be considered valid only if the following two conditions are satisfied: (1) the energy of differential rotation must be replenished very frequently, for all practical purposes continuously; and, (2) the density of this energy should be greater than the density of magnetic energy in fluctuations.

The first condition is associated with the fact that the excess of the energy of differential rotation is small, and this energy if not supplemented quite frequently will soon be exhausted, i.e., solar activity leads to rigid solar rotation. The second condition is also obvious: if the energy density of differential rotation is less than the magnetic energy in a fluctuation, it would be necessary to assume further the existence of a certain mechanism concentrating the energy of differential rotation, and the effect of this mechanism in turn would require the expenditure of energy.

We will now consider to what degree the first condition is satisfied, that is, whether there is a sufficiently frequent replenishment of the energy of differential rotation. We will not, however, discuss the mechanism of this replenishment. The order of time during which differential rotation should be changed into rigid rotation under the influence of forces of viscosity is given, according to Cowling (Ref. 8), by the following expression:

Table 24

Epoch of maximum of fluctuation	Energy of fluctuation, in ergs
August, 1917	$7.1 \cdot 10^{34}$
December, 1917	7.8
July, 1918	2.0
June, 1919	2.4
March, 1920	1.7
July, 1921	0.5

$$t = \frac{\rho L^2}{\mu}; \tag{3.3}$$

taking into account that the coefficient of dynamic viscosity $\mu = K\rho$ and the coefficient of kinematic turbulent viscosity $K = vL$, we have

$$t = \frac{L}{v}. \tag{3.4}$$

With the characteristic scale $L = 10^{11}$ cm, and the characteristic velocity $v = 10^2$ cm$\cdot$sec^{-1} (the velocity of general circulation of the sun) we obtain $t = 10^9$ sec, i.e., about 33 years (Ref. 8). However, it would be more correct to take not the velocity of general circulation, but the linear velocity of rotation, which considerably exceeds the velocity of circulation. When $r = 6.5 \cdot 10^{10}$ cm (convective zone) and $\omega = 3 \cdot 10^{-6}$ sec^{-1}, we obtain $2 \cdot 10^5$ cm/sec. With the same characteristic scale, this gives $t = 10^6$ sec, or about 10 days. We note that this value should be regarded as an upper limit since the characteristic scale $L = 10^{11}$ cm very probably is too high. It should be remembered that such a scale applies only in a horizontal direction, and not to the vertical direction, constituting only a very small part of the above-cited value. Thus, under conditions of turbulent viscosity, the energy of differential solar rotation should be replenished virtually continuously. From such a point of view this energy can serve as the energy source of solar activity fluctuations.

We will now consider the relationship between energy densities. We first will estimate the energy density of differential rotation. As already mentioned, for the inner zone of radiative equilibrium (that part of it where differential rotation still is expressed to some degree), the convective zone and the photosphere, the energy of differential is $6.0 \cdot 10^{32}$, $8.2 \cdot 10^{35}$ and $7.9 \cdot 10^{33}$ ergs, respectively. The volumes of the corresponding layers are: $3 \cdot 10^{32}$, $6 \cdot 10^{32}$ and $4 \cdot 10^{32}$ cm^3, respectively. Therefore, for the energy densities of differential rotation we will have:

In the inner zone of radiative equilibrium—2 ergs$\cdot$cm^{-3}.

In the convective zone—$1.4 \cdot 10^3$ ergs$\cdot$cm^{-3}.

Photospheric layers—13 ergs$\cdot$cm^3.

Thus, the more or less considerable energy density of differential rotation occurs only in the convective zone. A computation of the energy density of fluctuations made on the basis of data supplied by R. N. Ikhsanov gives $6.4 \cdot 10^5$ ergs$\cdot$cm^{-3}. It follows, therefore, that the energy of differential rotation cannot be the source of energy of solar fluctuations.

We will now compare the magnitude and energy density of differential rotation and general solar circulation. In accordance with the conclusions drawn in the preceding chapter, we have the right to assume that general circulation with a velocity of 10^2 cm$\cdot$sec^{-1} occurs in the layers to the base of the convective zone; in the deeper layers the circulation occurs at a far lesser velocity, approaching the velocity of Eddington currents, i.e., constituting not more than 10^{-4} cm$\cdot$sec^{-1}.

If it is assumed that circulation with low velocities extends to the depth for which $\frac{r}{R} = 0.30$, where r is the radius vector of the base of the layer of a particular depth and R is solar radius, the mass included between the base of the convective zone and the particular depth will be 0.372 of the solar mass (Wyman model) or $7.44 \cdot 10^{32}$ g. If the velocity of circulation is 10^{-4}cm$\cdot$sec^{-1} the energy will be $3.72 \cdot 10^{24}$

ergs, at the same time that the energy of circulation in the convective zone and photosphere is altogether $5 \cdot 10^{34}$ ergs.

Thus, the contribution to the energy of circulation of those layers where circulation occurs at a far lesser velocity is not significant. The determined energy of solar circulation is comparable to the energy of fluctuations of solar activity, and is not less than that for the differential rotation of these same layers. The energy density of the circulation is 10^2 $\text{ergs} \cdot \text{cm}^{-3}$, which is less than the energy density of the differential rotation for these same layers and especially the energy density of fluctuations.

We have combined the results in Table 25 for greater clarity and for use in further discussion of the relationship between the following energy characteristics: (a) solar activity; (b) differential solar rotation; and, (c) general circulation on the sun. It follows from this table that if we suspect a causal relationship between solar activity, differential solar rotation and general circulation on the sun, the first energy source in this chain should be solar activity, followed by differential rotation, with general circulation last.

It must be emphasized that from a comparison of the energies of differential rotation and general circulation it follows that rotation is of a primary character and circulation is secondary. It is obvious that it cannot be concluded from the above that solar activity is the only source of the energy of differential solar rotation. As pointed out in Ref. 8, the following qualitative model can be postulated: magnetohydrodynamic waves associated with active regions convey not only radiation and convection, but also negative angular momentum from the deep layers.

As we have seen, however, the relaxation time of differential rotation is small, and in an epoch of minimum solar activity, solar rotation should become rigid since in such epochs the transfer of momentum by activity should attenuate appreciably. However, it cannot be denied that solar activity plays a certain, although secondary, role in maintaining differential rotation.

Section 3. Endogenic and Exogenic Sources of Energy of Solar Activity

The energy estimates made in the preceding section provide a certain orientation in attempts to establish the cause of the large-scale energy transformations on the sun. With respect to the innermost sources of energy of this activity, or to be more exact, the mechanism of transformation of primary nuclear energy into the energy of cyclic activity,

Table 25

Form of energy	Layer affected by phenomenon	Total quantity of energy, ergs	Energy density, ergs/cm^3
Magnetic energy of sunspots (as a whole)	—	10^{36} (year)	$3.6 \cdot 10^5$
Magnetic energy in mean fluctuation	—	$3.6 \cdot 10^{34}$	$6.4 \cdot 10^5$
Differential rotation, taking into account attenuation of the differential effect with depth	Upper part of inner zone of radiative equilibrium	$6.0 \cdot 10^{32}$	2
	Convective zone	$8.2 \cdot 10^{35}$	$1.4 \cdot 10^3$
	Photosphere	$7.9 \cdot 10^{33}$	13
General circulation	Inner zone of radiative equilibrium	$3.7 \cdot 10^{24}$	—
	Photosphere and convective zone	$5.0 \cdot 10^{34}$	10^2

this problem remains almost entirely unsolved. We will now consider the concept of torsional oscillations from the energy point of view: This concept has been proposed by various authors, beginning with Walen (Ref. 11).

This hypothesis was designed to explain the development of the toroidal field from which the sunspot fields are formed. It is thought that the zone between the active zones of both hemispheres of the sun varies and induces a toroidal field at its margins, i.e., precisely where the zones of spot formation appear at the appropriate time. The variations of the zone are associated with convection, and their period can be determined by the general circulation cycle (Ref: 12). During the time of his work on this problem Walen (Ref. 13) proceeded on the assumption that in the sun's convective core, convection occurs in bursts, the core pulsates, and as a result of conservation of angular momentum, a magnetohydrodynamic wave will develop.

A model which considers variations of the zone of activity is more up-to-date. The strength of the induced toroidal field can be computed using a formula similar to that used for determination of the velocity of a magnetohydrodynamic wave, but in this case the latter is considered known: it is assumed to equal the velocity of circulation.

The formula has the form:

$$h = v\sqrt{4\pi\rho}. \tag{3.5}$$

If v and ρ are replaced by their values here, and the value 10^3 cm·sec^{-1} is assumed for the first, which is approximately an order of magnitude greater than the velocity of general circulation on the sun and, therefore, too high, and a value of 10^{-4} is assumed for the second, which approximately corresponds to the base of the convective zone, we obtain h = 30 oe.

It is possible that for some reasons the strength of the toroidal field will be greater, but it appears completely improbable that the energy source for induction of the toroidal field is general circulation. This assertion can be made on the basis of the energy estimates made in the preceding section. It is possible that the induction energy does not come from the energy of general circulation: the latter can impart only its own periods, that is, act like a pendulum in a clock mechanism. In general, however, this problem still has been studied very little.

The existence of 22-year pulsations of the solar diameter (see Chapter 1) can be highly promising for the problem of the mechanism of the transfer of energy from nuclear reactions to solar activity. If it is assumed that the results obtained by the Italian investigators are reliable, the amplitude of such pulsations on the basis of unsmoothed values exceeds 1", that is, 725 km on a linear scale. However, with such a prolonged period of variation as 22 years, the linear velocity of moving masses is only 10^{-3} cm·sec^{-1}. With such a small velocity, the kinetic energy, even if it is assumed that the entire solar mass participates in the oscillatory motion, is only 10^{27} ergs, which is much less than the magnetic energy dissipated during the cycle. It therefore follows that without the theory of oscillatory processes of the sun, it is impossible to draw conclusions from the presence of 22-year pulsations (even if they actually exist) that would assist appreciably in solving the problem of the energy sources of solar activity.

We pointed out in Chapter 1 that in the numerous attempts to represent solar activity, and especially its 11-year cycle, by the superposing of a series of periodic curves, investigators usually have been guided by a working hypothesis based on the postulation of a periodic process on the sun. The physical bases of such a periodic process were discerned in the tidal effects of the planets on the sun. It was felt, in particular, that Jupiter could be responsible; its period of revolution was conveniently close to the 11-year solar cycle for these supporters of the tidal hypothesis. However, the very great number of statistical investigations in this field did not lead to definite results, although certain authors even felt that it was possible to predict solar activity on the basis of planetary constellations. Estimates of the quantitative values of tidal effects, that is, attempts to place the solution of the problem to some degree on a physicomathematical base, were given far less attention.

We will discuss in greater detail one of these attempts to estimate the tide-generating force of the planets on the sun (made in 1954 by Anderson, and already discussed in Chapter 1 in connection with a 15-cycle period (Ref. 14)). Anderson considered that the relative tidal forces of the planets, beginning with Mercury and ending with Saturn, acted on the sun. The following is the computation method, using the example of Jupiter: the mean distance of Jupiter from the sun is 778,129,000 km. Since the radius of the sun is 696,000 km, the distance of Jupiter from the near and far sides of the sun is 778,825,000 and 777,434,000 km, respectively; the square of the ratio of these distances is 1.00356, that is, the relative difference of the forces of attraction on the near and far sides of the sun is 0.00356. Expressing the mass of Jupiter in earth units, we have 316.94, and expressing the distance to Jupiter in astronomical units, we have 5.2, hence $M/D^2 = 11.85$. After multiplying this by 0.00356, we obtain 0.0422. This is the relative tidal force of Jupiter on the sun. It is also important to estimate the ratio of the tidal force to solar attraction. The approximate formula for this computation has the form

$$\frac{G'}{G''}=\frac{2m'}{m''}\left(\frac{r}{R}\right)^3. \tag{3.6}$$

Here G' is the tidal force; G" is the gravitational force of the sun; m' is the mass of the tide-generating body; m" is the mass of the body on which the tide develops, in this case the sun; r is the solar radius; and, R is the distance between the celestial bodies. Since the mass of Jupiter relative to the sun is 0.001, and the ratio r/R in this case also is close to 0.001, formula (3.6) gives

$$\frac{G'}{G''}=2.10^{-12}.$$

This is 1/50,000th of the tidal influence of the moon on the earth; however, taking into account that the acceleration of gravity on the sun is 27.4 times greater than on the earth, the effect will be more significant. As a result, the tidal effect of Jupiter on the sun is 1/2,000th of the effect of the moon on the earth. Anderson computed that under the most favorable conditions, with the joint effect of all the planets on the sun when all were arranged in a single line relative to the sun, their joint tidal effect on the sun in relation to solar attraction would be $5 \cdot 10^{-12}$, which of course is completely insignificant. In the convective zone, however, where there is gravitational instability, the effect of the tidal influence can be more significant. A detailed consideration of the problem of the manifestation of the tidal effect of the planets in the zone of gravitational instability is beyond the scope of this book.

We have discussed this problem in considerable detail because investigations of the tidal influence of the planets on the sun are still appearing, such as a number of papers by Link (Ref. 15). These studies have a purely empirical-statistical character and warrant the criticism given them. Theoretical estimates of the possible effect are, in our opinion, far more useful than the numerous, but unpromising empirical attempts to detect such tidal influences.

In discussing the exogenic factors involved in solar activity, we of course should mention for the sake of completeness the meteorite hypothesis which was once popular. Currently, however, this hypothesis, to the effect that the energy of solar activity is supplemented by meteor matter falling into the solar atmosphere, is of purely historical interest.

Section 4. Estimating the Energy of Active Solar Radiation

The results presented in the preceding sections of this chapter, and the concepts presented there, show that at the present time the mechanism for transformation of the thermal energy released in the sun's interior into the energy of cyclic activity is still unclear. The numerical computations in Section 1 likewise should not give the idea that the total quantity of the energy of solar activity is already known.

It is obvious that the magnetic energy dissipating during the decay of sunspots is only a part of the total energy of activity. In addition, if we accept Cowling's concept that the magnetic fields of sunspots are not destroyed, but sink into the deep subphotospheric layers, it is no longer easy to insist on an unambiguous relationship between the magnetic energy of spots and the energy of cyclic activity. However, Cowling's contention does not appear to be sufficiently substantiated.

The magnetic energy, nevertheless, constitutes a value more or less actually known to us, and it is therefore natural that almost all energy estimates involved in this problem begin with its calculation. It is of interest to compare it with the total energy of those forms of solar radiation which are the result of its activity and either are absent in the quiet sun or have, during this state, a definite standard value characterizing some "background".

A quantitative estimate of active solar radiation involves great difficulties. Whereas the total quantity of radiant energy emanating from the sun can be estimated without difficulty, by proceeding on the basis of the well-known fact that the earth intercepts $\frac{1}{2} \cdot 10^{-9}$ of this energy, it is a different story with the radiation of the active sun. If we are concerned with wave radiation, it must be taken into account that it is associated primarily with a particular zone of activity. Furthermore, within this zone active wave radiation originates from discrete sources of definite dimensions, first appearing and then disappearing.

The problem is more complex in the case of corpuscular radiation. It is necessary to calculate the solid angles within which corpuscular streams are propagated. As is known, these angles are different for the slower particles causing magnetic storms with a gradual commencement, and having a 27-day cycle, than for the more rapid corpuscles associated with flares. Despite these difficulties, even now it is possible, on the basis of the abundant data obtained in recent years, especially during the IGY and IGC periods, and mostly by means of extra-atmospheric measurements, to form a certain idea concerning the energy value of active radiation.

This entire group of problems was recently generalized successfully by Biermann and Lüst (Ref. 1). We will cite here the concepts developed by these investigators, although only briefly. Biermann and Lüst classify as "nonthermal" those phenomena which appear to be the result of the transfer of energy in certain subphotospheric zones, or partly in the photosphere (atmosphere) itself, of a star by mechanisms other than radiation. It is not mandatory that such mechanisms transfer all the energy; it is sufficient that only a part be so transported. These phenomena are associated closely with hydromagnetic processes in the interior of a star, in our case the sun. Four types of solar active radiation are known (and only certain of them are truly nonthermal):

(1) excess electromagnetic (i.e., wave) radiation, exceeding the level corresponding to the temperature of the photospheric layers in the Lyman and X-radiation regions;

(2) excess electromagnetic radiation in the radio range, observed below the ionosphere, from 15-20 to 50,000 Mc/s;

(3) corpuscular radiation, usually in the form of streams of ionized gas;

(4) corpuscular radiation, sometimes in the form of relativistic particles with energies from 30 to 100 Mev and up to 50 Bev.

As is well known, the flux of thermal energy is of the order of $6 \cdot 10^{10}$ ergs $cm^{-2} \cdot sec^{-1}$. The ultraviolet radiation in the Lyman region comes (for the most part) from the chromosphere, and in the X-radiation region from the corona. According to rocket data, in the Ly_{α} flux the energy ranges from 10^3-10^5 to 10^6 ergs$\cdot cm^{-2} \cdot sec^{-1}$ on the solar surface. The intensity of X-radiation can be determined both from the temperature and density in the corona and by rocket methods.

Both methods give results in good agreement—the energy flux is of the order of $5 \cdot 10^3$ ergs$\cdot cm^{-2} \cdot sec^{-1}$. The energy radiated at the radio frequencies from the chromosphere (frequency $\nu = 10{,}000$ Mc/s) under undisturbed conditions is 10^{-1} erg$\cdot cm^{-2} \cdot sec^{-1}$; however, at the radio frequencies with a lesser frequency ($\nu = 100$ Mc/s) emitted by the corona, it is 10^{-4} erg$\cdot cm^{-2} \cdot sec^{-1}$. Long-lived active regions yield fluxes exceeding by two orders of magnitude those indicated above.

We will now consider corpuscular radiation. On the basis of geomagnetic observations and investigations of type-I cometary tails it can be concluded that there is an almost constant presence of certain corpuscular radiation. The mean flux of its energy is 10^{-5}-10^{-6} erg$\cdot cm^{-2} \cdot sec^{-1}$. At a distance of one astronomical unit this gives 10^9 ions$\cdot cm^{-2} \cdot sec^{-1}$. At the time of solar flares there is the strongest deviation from thermal equilibrium. The total quantity of nonthermal energy emitted by powerful flares can attain $\approx 10^{33}$ ergs in 10^3 seconds. At the time of the large flare of February 23, 1956, the nonthermal radiation in the form of cosmic radiation was 10^{31} ergs. In general, at the time of large flares the maximum energy of particles can be from 10 to 50 Bev, but there are events when the maximum energy of the corpuscular stream attains only 100 Mev or even less.

It should be noted that the total flux of radio emission at the time of the flare of February 23, 1956, was 10^{25} ergs. When there is a less significant production of relativistic particles their total energy can be 10^{27}-10^{28} ergs, which nevertheless is two or three orders of magnitude greater than the energy emitted in the radio range, even at the time of the strongest events. Biermann and Lüst draw an important conclusion: in the case of a stellar structure the transition of from 0.01 to 1 percent of the thermal energy into "nonthermal" (that is, active) radiation of some kind (wave or corpuscular) should be quite expected. Table 26 gives the values of solar active radiation for different conditions.

As shown by observations of certain stars, especially stars with "flares", the type of activity observed on the sun may not be a rare phenomenon, but may be entirely normal for stars of the late type. From this point of view, the sun should be considered a star of low activity. Biermann and Lüst consider the convective zone to be responsible for the development of active radiation. The direct mechanism responsible involves acoustic and magnetohydrodynamic waves arising in the convective zone and propagating into the higher layers.

The low activity of the sun as a star forces us to assume that the fraction of thermal energy passing into active radiation is closer to the lower limit given by Biermann and Lüst than to the upper limit. If it is assumed that this fraction is 0.01 percent of the thermal radiation of the sun, the sun will release 10^{29} ergs in 1 second in the form of active radiation. This will be $3 \cdot 10^{36}$ ergs annually. This value is quite close to the magnetic energy of sunspots expended in a year, and if we also take into account the energy of weak magnetic fields, it will be admissible to assume that the dissipation of magnetic fields is the source of active solar radiation.

Variations of geoactive solar radiation within the limits of the 11-year cycle, when there is virtually total constancy of the solar constant, force us to assume that within the limits of the cycle there is some change in the fraction of energy emitted by means of "nonthermal" mechanisms. In the epoch of a maximum this fraction will be somewhat greater, and in years of the minimum, it will be somewhat less. At the present time systematic reliable extra-atmospheric investigations do not cover even a single full 11-year cycle. It would therefore be somewhat premature to raise the problem now of the cyclic amplitude of any form of radiation from among those listed in Table 26.

Table 26

Type of activity	Electromagnetic radiation		Corpuscles	
	excess at Lyman and X-radiation frequencies	excess at radio frequencies from 15-20 to 50,000 Mc/s	fluxes of ionized matter	relativistic particles from 30-100 Mev to 50 Bev
Mean, in absence of special activity (in dependence on phase of cycle).	10^5-10^6	10^{-1}	—	—
Long-lived active regions. .	10^6	10^0-10^1	—	—
Flares	10^8-10^{10}	10^2-10^3	10^7-10^9	10^7-10^8

Note: all values are in $ergs \cdot cm^{-2} \cdot sec^{-1}$ on the surface of the sun

CHAPTER 4
CERTAIN MANIFESTATIONS OF SOLAR ACTIVITY IN THE UPPER LAYERS OF THE EARTH'S ATMOSPHERE

Section 1. General Remarks

There is an exceptional variety of manifestations of solar activity in the upper layers of the earth's atmosphere. The relationships between solar activity and processes in these layers are extraordinarily close and it can be stated without equivocation that the region of the upper layers of the earth's atmosphere is determined by solar activity. No matter whether we deal with the state of the ionosphere, various kinds of radiation of the upper layers (auroras, night airglow), or the state of the geomagnetic field, in all cases solar radiation must be taken into account; this solar radiation induces a number of processes in the upper layers of the earth's atmosphere. This radiation is expended in inducing and maintaining these processes and does not reach the earth's surface, nor the lower layers of its atmosphere. Certain authors assume that the upper layers of the earth's atmosphere can be considered as a sort of continuation of the outermost layers of the solar atmosphere. For example, Chapman has postulated that the outermost part of the solar corona extends to the earth's orbit (Ref. 1).

Investigations of the upper layers of the atmosphere and space by rockets and artificial satellites have greatly supplemented our information on the conditions prevailing in these layers and the physical processes transpiring there. As a result of these investigations, in a number of cases it has been necessary to discard accepted concepts and the earlier theoretical interpretations of observed phenomena, and to develop new ones. We will cite certain examples: whereas it was assumed earlier that the processes in the upper layers are caused directly by particular forms of solar radiation, i.e., that this radiation is primary for the generation of such processes, it now is considered possible that secondary radiations develop in the atmosphere itself under the influence of primary radiation. These secondary processes exert an ionizing effect. Whereas formerly it was assumed that corpuscular streams have no stably existing condensations in the immediate neighborhood of the earth, the discovery of the radiation belts has forced abandonment of this point of view.

It should be added at this point that advances in the field of magnetohydrodynamics and gasdynamics have made possible new explanations for certain phenomena, such as the sudden commencement of magnetic storms now associated with waves in extended plasma of solar origin. The abundance of new facts and the new possibilities of interpretation of observed facts have seriously shaken old concepts concerning many processes in the upper layers of the atmosphere.

Whereas only relatively recently there was a more or less generally accepted opinion concerning how solar radiation is responsible for the ionization of the various layers of the ionosphere, there now are great possibilities for formulation of appropriate hypotheses, and this has led to less certainty in solving the problems involved. Whereas in former years most investigators assumed, although with reservations, that the Chapman-Ferraro theory of magnetic storms was correct, and most work was on the perfection of individual aspects of the mechanism proposed by these authors, a number of new hypotheses now have been advanced and quite some time must now pass before it will be possible to settle on anything definite.

The manifestation of solar activity in the upper atmosphere has always been considered extremely extensive, and rightly so. If we also take into account that "overestimate of values" which at present is so characteristic of this branch of science, it becomes completely obvious that it is impossible to discuss in this monograph all or even a considerable part of the problems involved. With this in mind, we will concentrate attention on only a single problem: the heating of the upper layers of the atmosphere by active solar radiation. It is not without a reason that this problem is selected. This problem is related directly to the problem of the mechanism of the effect of solar activity on phenomena in the lower layers of the earth's atmosphere; the latter problem will be the subject of Chapter 5.

Section 2. Observational Data on the Relationship of Ionospheric Temperatures to Solar Activity

A certain amount of data on the temperature in various layers of the ionosphere has already been accumulated. Unfortunately, if we are concerned with the study of relationships with solar activity, in the sense of a relationship with the phase or particularly the number of an 11-year cycle, these data are still inadequate. A comparison for lesser intervals of time is possible.

Despite the fact that as a result of extra-atmospheric flights it has proven possible to determine the temperatures of the upper layers by direct methods, for the time being we can obtain long series of values only by indirect methods. There is a considerable variety of such

methods. In this section we will discuss only radio methods, since the extensive network of ionospheric stations existing at the present time provides considerably more data than optical and other methods.

With respect to purely ionospheric methods, we should mention three: frequency of collisions, effective recombination coefficient, and height of the homogeneous atmosphere. The first method—use of the frequency of collisions—is based on completely reliable relations between the number of collisions of electrons with neutral particles, electrons, and ions and temperature. In most cases the method involves determination of the effective number of collisions in which the numbers of collisions with different types of particles are included as terms. The principal difficulty in determination of temperature by this method is finding the numbers of collisions, since this requires special observations not included in the program of most ionospheric services. The collision frequencies are determined either by the use of cross modulation of radio waves from two radio stations or from measurement of the coefficient of reflection from the corresponding ionospheric layer, as a function of the incident frequency. Rather numerous determinations of the frequency of ionospheric collisions have appeared recently, but for the time being the data obtained by this method is inadequate for comparison with solar activity.

More extensive data have been obtained by the method of determination of the effective recombination coefficient. The essence of this method is as follows. The ionization-recombination process, as is well known, is described by the equation

$$\frac{dN}{dt}=I-\alpha N^2, \tag{4.1}$$

here N is the electron concentration; t is time; I is the intensity of ionizing radiation; and, α is the recombination coefficient. The value α changes during the course of the day, but for two times symmetric to true local midday, both values α should be equal to one another. Obviously, the values I also are equal to one another for two such times. After denoting the electron concentrations for two such times, symmetric to midday, by N_1 and N_2, respectively and their changes with time near the corresponding hours by $\left(\frac{dN}{dt}\right)_1$ and $\left(\frac{dN}{dt}\right)_2$, we obtain from formula (4.1) for the two times

$$\left(\frac{dN}{dt}\right)_1=I-\alpha N_1^2,$$

$$\left(\frac{dN}{dt}\right)_2 = I - \alpha N_2^2,$$

which gives

$$\alpha = \frac{\left(\frac{dN}{dt}\right)_1 - \left(\frac{dN}{dt}\right)_2}{N_2^2 - N_1^2}. \tag{4.2}$$

It is also possible to find the recombination coefficient from data on nighttime ionization. In the absence of ionizing radiation, that is, when there are only recombination processes, equation (4.1) assumes the form

$$\frac{dN}{dt} = -\alpha N^2, \tag{4.3}$$

and it is then possible to determine α. It must be remembered that this latter relation can be used only infrequently. For example, the fact is that in the case of the E layer nighttime ionization often has very small values, not making it possible to record any definite critical frequency. Therefore, in most cases it is necessary to use a formula for near-midday values.

It is obvious that when there is an equality of electron concentrations (and therefore of critical frequencies) equation (4.2) no longer applies, but in many cases N_1 and N_2 may not be equal, but are close to one another. The same can be said relative to the values $\left(\frac{dN}{dt}\right)_1$ and $\left(\frac{dN}{dt}\right)_2$. Computations made by the use of formula (4.2) can be burdened by purely computable errors. In addition to inaccuracies associated with determination of the recombination coefficient, this method of determination of ionospheric temperatures has still another shortcoming associated with the known ambiguity in the relationship between the recombination coefficient and temperature. The relationship between T and α usually is given in the form established by Thompson, that is

$$\alpha = CT^{-n}, \tag{4.4}$$

where n and C are constants; hence

$$T = T_0 \sqrt[n]{\frac{\alpha_0}{\alpha}}; \tag{4.5}$$

here T_0 and α_0 are the temperature and recombination coefficient for certain initial conditions. In the first approximation it is assumed that n = 3. The selection of the initial values of temperature and the recombination coefficient to a certain degree is arbitrary, which inevitably exerts an influence on the values of the determined ionospheric temperatures. This is of course another shortcoming of the method. Then, for the F_2 layer the simple ionization-recombination equation (4.1) does not always adequately describe the processes occurring there. V. P. Kolokolov (Ref. 2) has replaced it with the following:

$$\frac{dN}{dt} = I - \alpha N^2 - \frac{N}{T}\frac{dT}{dt}. \tag{4.6}$$

The last term should take into account the thermal expansion of the layer. However, the theory of the F_2 layer has not yet been developed adequately, and it is unclear whether the last term in (4.6) takes such processes into account. Kolokolov assumes that for the E and F_1 layers and for the F_2 layer in winter it is possible to use the simple equation (4.1). In his study Kolokolov used T = 230°K as the initial temperature for the E layer. This value was taken from spectral observations of auroras at Tromso. In the case of daytime conditions the initial temperature used was T = 330°—a value obtained using data from rocket ascents in New Mexico (United States). For the initial value of the recombination coefficient in the E layer for daytime hours the value used was $5\cdot10^{-9}$ cm$^3\cdot$sec^{-1} (data for the ionospheric station at White Sands, New Mexico). The initial value used for nighttime was the value α_0 in the first morning hours for the corresponding station. With respect to the F layer, α_0 was considered to equal the mean nighttime value for Central Asia—$5\cdot10^{-10}$cm$^3\cdot$sec^{-1}. The initial temperature used for this layer was obtained from twilight observations at the Abastumani Observatory (Ref. 3), i.e., 750°K. The mere listing of these values shows how diverse they are with respect to the place where they were determined and the method by which they were obtained. This inevitably had an effect on the results and also is evidence of inadequacy of the method.

The ionospheric data in Kolokolov's study was obtained from many stations: Watheroo, Ottawa, Trinidad, etc. As a result of the investigation, the author obtained a number of dependencies of temperatures of ionospheric layers on geographic latitude, season, etc. We are concerned

here with the relationship between these temperatures and solar activity. Figure 36 shows a curve illustrating such a relationship on the basis of data for January and February for Watheroo station for 1938-1947, and for the F layer. This figure, taken from the Kolokolov study, shows that in years of high solar activity the temperature of the layer is approximately 150° higher than in the epoch of a minimum. Kolokolov also points out an increase of the temperature of the E layer from the minimum of the solar cycle to its maximum.

The third method for determination of ionospheric temperatures is based on the relationship between critical frequencies and the height of the homogeneous atmosphere, and the relationship of the latter to temperature. The relationship between electron concentration and the height of the homogeneous atmosphere H is different for different models of the layer (simple, parabolic, etc.). For example, the following method can be used (Ref. 4): the minimum virtual height of this layer at midday of a particular day and the critical frequency for this same time are taken from ionospheric tables for the F_2 layer.

By multiplying the value of the time by 0.834 (Ref. 5) and entering it in the height-frequency diagram, it is possible to obtain the height of maximum ionization. The difference between this value and the minimum virtual height gives double the height of the homogeneous atmosphere. But this method is not only very approximate, but also in many cases it is simply useless, as there are cases in which it is not possible to obtain from the standard height-frequency diagram the height of maximum ionization for a frequency equal to 0.834 of the critical frequency.

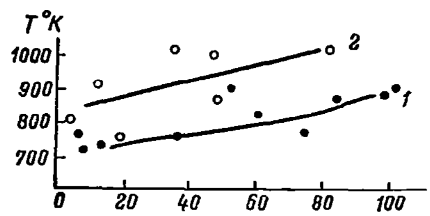


Figure 36. Dependence of January (1) and February (2) temperatures of the F_2 layer for Watheroo station on

Wolf number (horizontal axis) for 1938-1947 (according to V. P. Kolokolov)

The most important shortcoming of the method involving determination of the height of the homogeneous atmosphere is that temperature cannot be determined with the required accuracy, due to a lack of knowledge concerning the precise molecular weight. The mean molecular weight of the air at a particular height can differ appreciably from its value on each specific day. It is obvious that appreciable changes of molecular weight can occur during short intervals of time. When using the mean molecular weight, we either underestimate or exaggerate the temperature determined from the relation

$$T = \frac{\mu g}{\Re} H, \tag{4.7}$$

where μ is the molecular weight of air; $\Re$ is the universal gas constant; and, g is the acceleration of gravity.

Variations in the molecular weight of air can be caused either by an intensification of dissociation of some gas (in the case of the F_2 layer this would be the dissociation of molecular nitrogen), or intensified vertical exchange with lower-lying layers, or, since we are concerned with local changes μ, with some horizontal mixing of air bringing into this region air masses with a different molecular weight.

It will be demonstrated below that when data are averaged for a period of approximately one month the mean molecular weight for a particular height already is quite close to what could be called its "climatic value". With respect to determination of the height H of the homogeneous atmosphere itself, certain observational data make it possible to avoid difficulties associated with computations of the height of the ionization maximum, since such data directly contain the half-thickness of the layer.

An example of such data are the observations of the Japanese ionospheric station Kokubunji in Tokyo. The height of the homogeneous atmosphere is determined for the applicable layer model as half the half-thickness. The molecular weight and the acceleration of gravity are taken from a table for the height of the ionization maximum, which also is given in observational tables. The results of investigations of the relationship between solar activity and the temperatures of the F_2 layer, determined by the method described, are as follows.

The first step was to determine active regions, as defined by D'Azambuja, with an importance of $\geqslant 3$ (on a 10-unit scale) for 1951, which was one of the years on the descending branch of cycle No. 18 of solar activity. The year 1951 was very convenient in this respect because at that time active regions already were separated quite clearly from one another, but at the same time there were still a sufficient

number on the solar disk. The passage of these regions across the sun's central meridian was compared with the temperatures of the F_2 layer determined from the height of the homogeneous atmosphere. The comparison was made using the method of superposing epochs.

A total of 57 cases of the passage of active regions across the central meridian were considered. The computation of H was by the first method, i.e., from the midday critical frequency for a particular day. From the height-frequency diagram it was possible to determine the height corresponding to a frequency equal to 0.834 of the critical frequency. The minimum virtual height was determined from another table, and the half-thickness of the layer then was determined. Table 27 gives the temperature deviations from the corresponding monthly median values, averaged by the method of superposing of epochs, for different days prior to and subsequent to the reference day, and also for the latter, which corresponds to the passage of an active region across the central meridian.

The last line gives the number of active regions used in computations for the corresponding day prior to and subsequent to the reference day. These numbers are not equal to one another, and are not equal to the total number of 57 active regions in that in certain cases, as noted above, the height of the homogeneous atmosphere could not be obtained from the value of the critical frequency.

Table 27 shows that two days before the passage of an active region across the central meridian, there is an appreciable positive temperature anomaly. The probability of randomness of such a deviation, computed using the t test, is 0.013. It is true that in the computation of this probability the number of degrees of freedom probably was somewhat too high as it was impossible to take into account inertia in conservation of the temperature value. The results, with respect to a

Table 27

	Day from reference day												
	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6
Temperature deviations. . . .	12°	4°	8°	3°	27°	-8°	6°	-4°	5°	4°	-4°	13°	16°
Number of active regions	30	32	31	27	35	32	34	35	31	32	34	29	30

temperature increase in the F_2 layer at the time of passage of an active region across the sun's disk, are confirmed on the basis of data for 1961, although to be sure, only in a single very sharply expressed case of the passage of an isolated, quite intense active region with numerous flare centers. This computation was based on data for the Kokubunji station.

Figure 37 shows a synoptic map of the sun with spot groups of different area gradations (in millionths of a hemisphere) and flares; under the map there is a curve of variation of the temperature deviation of the F_2 layer for midday values at Kokubunji from the ~~median~~ monthly value for this same hour. The map and curve are for February, 1961.

About February 9, that is, one day before the passage of the center of the active region across the sun's central meridian, there was a positive temperature anomaly in the F_2 layer attaining 300° according to the smoothed values shown in the figure (unsmoothed data show that the value of this anomaly was still greater) (Ref. 6).

Thereafter an attempt was made to carry out a more detailed investigation of the relationship between the temperature in the F_2 layer and solar activity. To do this the data for the station Kokubunji for January, 1957-August, 1961, inclusive were used to find the mean annual temperatures of this same layer by the same method. They, as well as the mean annual values of the height of the homogeneous atmosphere, were compared with the mean annual Wolf numbers for these years. The results have been shown in Figure 38, which shows the clearly expressed parallel variation of Wolf numbers and the mentioned ionospheric characteristics.

The correlation coefficients between the mean monthly Wolf numbers and the median monthly heights of the homogeneous atmosphere, and also the mean monthly temperatures at the level of the F_2 layer, were computed. They were $+0.78\pm0.05$ and $+0.75\pm0.06$, respectively (the errors are mean values, the number of correlated pairs is $n = 53$).

Section 3. Theoretical Considerations Concerning the Relationship Between Ionospheric Temperatures and Solar Activity

As pointed out in the last section of Chapter 3, active solar radiation can be represented rationally as the sum of the corresponding radiation of the active sun and a certain "background" characteristic of the "quiet" sun. The undisturbed ionosphere is created by the shortwave

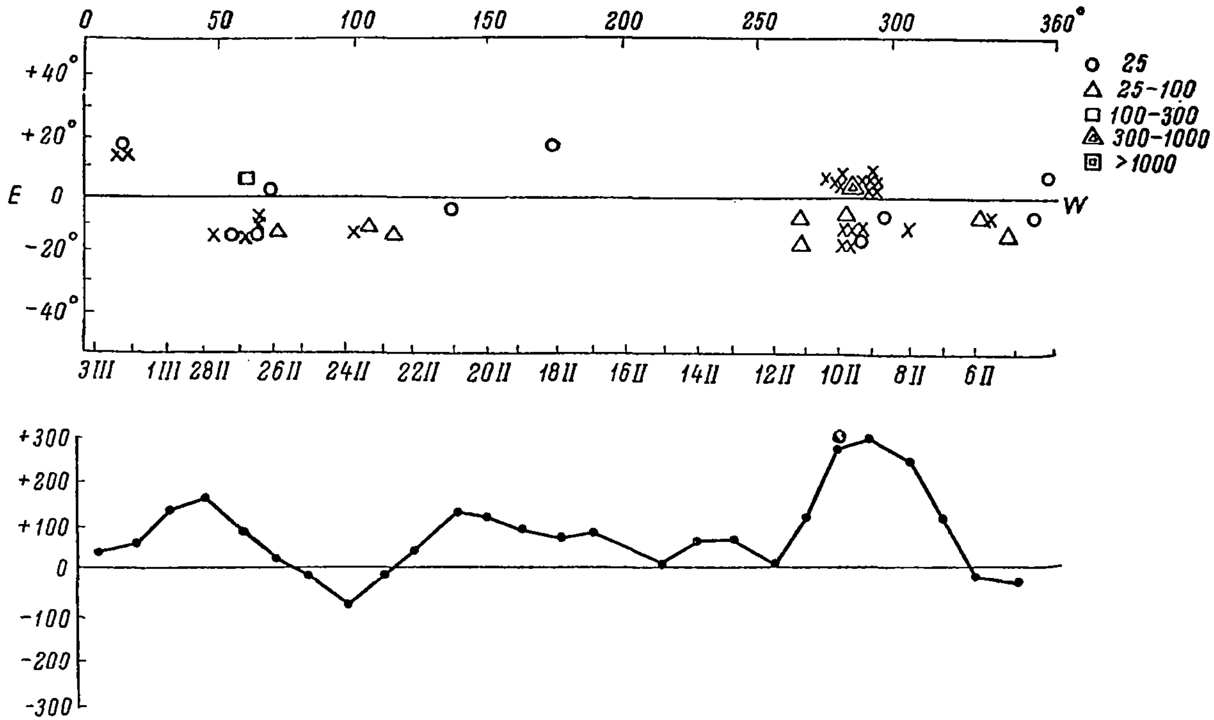


Figure 37. Synoptic map of the sun for February, 1961, and deviations of temperature of the F_2 layer from the monthly median. Flares have been denoted by crosses; the gradation of spot areas in millionths of a hemisphere is given in the figure

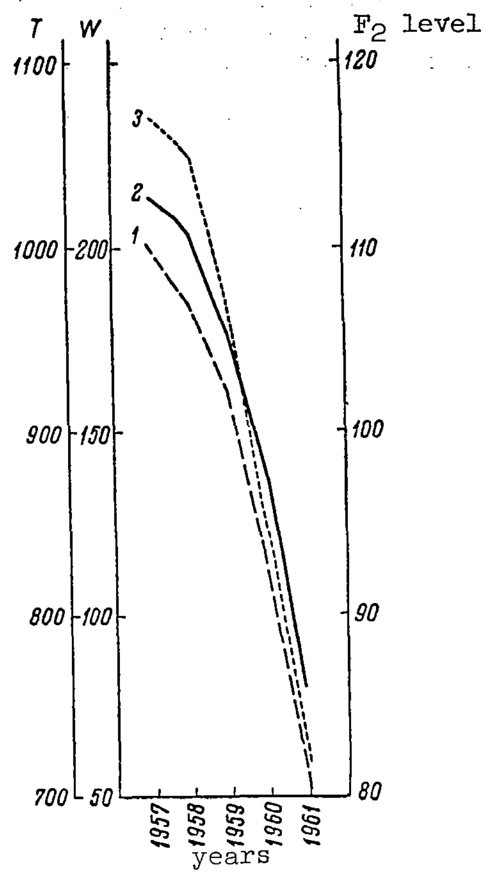


Figure 38. Ionospheric temperature and solar activity: 1, Wolf numbers; 2, mean annual temperatures of the F_2 layer for 1957-1961 at Kokubunji station (Japan); 3, mean annual doubled height of the homogeneous atmosphere at this same level

radiation of the "quiet" sun. Since the principal reason for the high ionospheric temperatures is ionospheric absorption of solar radiation, accompanied by ionization, there should exist a relationship between the temperature of the ionosphere and the electron concentration in it. This relationship can be determined rather easily from general theoretical considerations (Ref. 7).

Since solar wave radiation is responsible for ionization of the ionosphere and an increase of ionization with an intensification of solar activity, intensifications of solar wave radiation should lead to

an increase of temperature in the high layers of the atmosphere. The results cited in the preceding section confirm this conclusion. However, opinions diverge when there is a more thorough consideration of the relationship between solar activity and ionospheric temperatures.

At one time, particular attention was given to the D region of the ionosphere due to its most direct relationship to solar active wave radiation. We refer to sudden ionospheric disturbances expressed in sudden and strong increases of ionization in the D region. Only recently these phenomena were related to a considerable increase in the intensity of the L_{α} line in the flare spectrum in comparison with the intensity of this line in the spectrum of the undisturbed sun.

Prior to ascents of rockets into the upper layers of the atmosphere it was assumed that the intensity of radiation at Lyman frequencies inevitably should intensify, since at the time of a flare there is a considerable increase of emission at Balmer frequencies. Proceeding on this basis, it was assumed that the temperature in the D layer should increase appreciably at the time of a flare. For example, in 1950, Stranz (Ref. 8) calculated that the maximum increase of temperature at heights of 95-100 km at the time of a strong flare is about 30°/per half hour, that is, as a mean for the time of a flare. However, rocket ascents did not confirm the assumption of a considerably greater intensity of Lyman lines at the time of a flare in comparison with its intensity in ordinary undisturbed periods (Ref. 9).

A new theory was therefore developed concerning the origin of sudden ionospheric disturbances. It was postulated that they are caused by extremely shortwave radiation, i.e., by hard X-rays (Ref. 10). But although solar radiation also increases at such frequencies at the time of flares, its flux nevertheless is very small for an appreciable increase of temperature in the D region. Secondary radiation also must be taken into account, since it complicates the problem considerably. We therefore feel that it is premature to develop a theory of the change of the temperature of the D region at the time of these disturbances considering the present-day availability of information concerning the nature of the solar radiation causing sudden ionospheric disturbances.

With respect to the E and F_2 layers, the situation can be considered somewhat more favorable. Although with respect to the E layer it is impossible to say at the present time exactly what photon radiation of the sun actually is responsible for its ionization (it is most probable that this also is X-radiation, but softer), it can be concluded that there are variations of this radiation on the basis of ionization changes in the 11-year cycle. The equivalent photon emission, causing ionization in the E layer, increases from the epoch when the Wolf number

is equal to zero to the epoch when it is 100 by a factor of 2 (from $2.3 \cdot 10^{13}$ to $4.6 \cdot 10^{13}$ photons/$cm^2 \cdot sec^1$). For the F_2 layer this value increases by a factor of almost 3: from $1.1 \cdot 10^{14}$ to $3.2 \cdot 10^{14}$ photons/ $cm^2 \cdot sec^1$.

Therefore, if there is a heat balance equation for the corresponding layer it also is possible to compute the change of mean temperature. However, since the determination of the photon flux itself is done on the basis of ionospheric measurements, there is no sense in using this method and it would be more correct to use one of the techniques for determination of temperature from direct ionospheric observations.

It must be assumed that in the future, when solar shortwave radiation is recorded regularly by extra-atmospheric and interplanetary observations, computations of variations of ionospheric temperatures, based on the writing and solution of heat balance equations for the ionosphere, will be entirely possible, and when compared with direct temperature soundings, will make it possible to obtain much important information in the field of ionospheric physics.

There are more definite and finalized theoretical conclusions concerning the heating of the ionosphere by corpuscular solar radiation. The following approaches to the problem can be noted:

(1) there is heating directly by corpuscular solar radiation (elastic impact);

(2) there is a dissipation of electric currents associated with various kinds of geomagnetic disturbances; in other words, the heating of the upper layers by Joule heat liberated by the currents associated with magnetic variations and especially with magnetic activity;

(3) there is heating of the upper layers of the atmosphere by the interplanetary gas, to be more precise, in accordance with Chapman (Ref. 1), that is, by the outermost regions of the solar corona.

We will discuss the first two mechanisms as the most directly associated with solar activity, although it is altogether likely that further investigations of the third mechanism will reveal its close relationship to solar activity. However, for now it is unclear how coronal inhomogeneities (such as condensations), systematically observed in the inner corona, can manifest themselves in regions so distant from the sun. Thus, we will concern ourselves here with the mechanisms of elastic collisions and heating by Joule heat. It should be noted that in a number of cases it is difficult to distinguish the effects of Joule

heating associated with elastic collision from the effects produced by Joule heat. There is, therefore, no complete assurance that in the solution of the corresponding equations the same value is not involved twice.

The problem of the heating of the upper layers of the atmosphere at the time of magnetic storms has been formulated and approximately solved by Chapman(Ref. 11, 1937). Still earlier, in 1918, Chapman considered the problem of the energy of magnetic storms and approximately computed the energy released during the decay of the joint field, i.e., the field of the storm and the earth's field, at the time of their interaction (Ref. 12).

In a study published in 1937, Chapman investigated the heating by Joule heat of a conduction current which at the time of a magnetic storm flows approximately at the level of the E layer. The computation was made for the auroral zone, where the conduction current should be maximal. In these computations Chapman made a number of simplifying assumptions: he considered a current directed along the magnetic lines of force of the earth's field, whereas in the auroral zone this current was directed normally to the lines of force. Chapman also neglected the mechanisms of heat loss, i.e., he ignored the increase of this loss associated with an increase of the temperature gradient. On the basis of such assumptions, Chapman concluded that at the time of a strong magnetic storm the temperature at a height of 100 km increases by approximately 18° per hour.

In 1952, I. S. Shklovskiy (Ref. 13) analyzed the possibility of heating of the upper layers of the atmosphere at the time of auroras as a result of elastic collisions with solar protons. It was found that at a height of about 100 km, each cubic centimeter of air acquires about 10^{-6} erg/sec by this process. Such a heat increment is very small, and cannot lead to an appreciable temperature increase.

We will now consider the mechanism of elastic collisions in greater detail. In the previously discussed investigation of I. S. Shklovskiy, it was assumed that all the primary energy of the stream of solar corpuscles is expended on these collisions. Such an assumption obviously will give the upper limit of the increase of the kinetic temperature of the air. However, in general we must subtract from the primary energy of the stream: (1) the energy expended on ionization; and, (2) the energy expended on the excitation of the electron levels of N_2 and O_2 (and in part, of the oxygen atom) since the energy expended on the excitation of the vibration-rotation levels of molecules is set free for all practical purposes by elastic collisions, and not by

luminescence. The reason for this is the very low probability of vibration-rotation transitions of the particular molecules.

According to Shklovskiy, in the case of the most probable mechanism of propagation of a stream of solar protons from the uppermost layers of the atmosphere to the level of the E layer, i.e., in the case of the charge exchange mechanism, about 50 percent of all the initial energy of the stream is expended on ionization. It would appear that in actuality this loss is somewhat less, since the secondary electrons appearing in the ionization process also transfer part of the excess energy to molecules by means of elastic collisions. This begins to occur when the energy of the electron, as a result of inelastic collisions, becomes less than 1.96 ev, the excitation potential of the red oxygen line.

But the fact that the greater part of the energy of the secondary electrons is expended on ionization and excitation, as a result of the smallness of their mass in comparison with the mass of molecules, makes this secondary heating mechanism ineffective. At the present time it is customary to assume that the ionization and excitation accompanying auroras are associated with electrons of solar origin with energies from 1 to 30 Kev. These new data indicate a far smaller expenditure of the energy of protons on ionization and excitation with subsequent luminescence than was assumed earlier. Fast solar electrons do not penetrate into layers lower than 100-300 km in the earth's low latitudes, but easily can penetrate to such heights in the polar regions.

Thus, it is in this sense that we must understand Paetzold's assertion (Ref. 14) that solar particles do not penetrate to levels lower than 170 km. It can be assumed that about 40 percent of the initial energy of the stream is expended directly on the heating of the upper layers of the atmosphere by elastic collisions. We note once again that in the light of new concepts concerning the development of auroras this percentage appears too low, since protons expend little of their energy in inelastic processes. Nevertheless, we will retain this minimum coefficient 0.4, proceeding on the assumption that once again in the light of new concepts, the concentration of protons can be somewhat less than was assumed earlier.

The expression for the energy flux of protons, expended in inelastic collisions, can be written as

$$\varepsilon_T = 0.4\frac{mv^3[H]}{2}; \tag{4.8}$$

here m is the mass of the proton; [H] is the number of protons per 1 cm^3 at the height of the heated layer; and, v is the velocity of solar protons.

In order to estimate the heating, it is necessary to specify values of velocity and concentration. It is rational to make computations for several variants of the velocity value, such as 1,000, 2,000 and 3,000 km/sec. With respect to the concentration of particles of solar origin at the time of magnetic storms, it will be rational to assume that it falls in the range of 10^2-10^3 particles/cm^3, depending on whether a moderately strong or a very strong storm is considered. In order to obtain the energy released in 1 cm^3, it is necessary to divide the results of the computation based on use of formula (4.8) by the height of the homogeneous atmosphere, which for this level can be assumed to be 10 km. Results of the computations are given in Table 28.

Thus, in the numerators we have the fluxes (with the dimensionality $ergs \cdot cm^{-2} \cdot sec^{-1}$) and in the denominators, the energy increments (with the dimensionality $ergs \cdot cm^{-3} \cdot sec^{-1}$).

It should be noted that, according to the recent work of Cole (Ref. 22), at the time of auroras there is a strong "inflating" of the atmosphere, especially effective, to be sure, at heights somewhat exceeding the level considered here, but also probably leading to a considerable increase in the height of the homogeneous atmosphere at this level as well. The values in the denominators in Table 28, therefore, may be somewhat too high. However, since hereafter in all cases we will use the upper limits for the heat loss, we will assume that failure to take into account an increase of the height of the homogeneous atmosphere will not be of too great a significance.

Another mechanism for the heating of the upper layers of the atmosphere at the time of magnetic storms is the mechanism of release of Joule heat. This is heat released by a conduction current in 1 cm^3 in 1 sec and is described by the formula

$$q = \frac{i^2}{\sigma}, \tag{4.9}$$

where q is Joule heat; i is current density in the circuit; and, σ is electrical conductivity.

For this computation we must know: (1) the intensity of the conduction current at the time of a magnetic storm; (2) the area of the cross section of the circuit through which the current flows; and, (3) the electrical conductivity at the level of the current at the time of the magnetic storm. The intensity of the conduction current is

Table 28

Proton velocity	Quantity of protons H			
	1	10	100	1000
1000	$\frac{0.3}{3.0\cdot 10^{-7}}$	$\frac{3.0}{3.0\cdot 10^{-6}}$	$\frac{30}{3.0\cdot 10^{-5}}$	$\frac{300}{3.0\cdot 10^{-4}}$
2000	$\frac{2.7}{2.7\cdot 10^{-6}}$	$\frac{27}{2.7\cdot 10^{-5}}$	$\frac{270}{2.7\cdot 10^{-4}}$	$\frac{2700}{2.7\cdot 10^{-3}}$
3000	$\frac{9.0}{9.0\cdot 10^{-6}}$	$\frac{90}{9.0\cdot 10^{-5}}$	$\frac{900}{9.0\cdot 10^{-4}}$	$\frac{9000}{9.0\cdot 10^{-3}}$

determined from geomagnetic measurements. The basic difficulty involved in such determinations is that two different values can be found from the same set of geomagnetic data for the time of a magnetic storm: the current intensity and the distance to the current-carrying circuit. Theoretically, this problem is solved by measurements at two or more stations.

In the computations cited below it was assumed that the current flows at the level of the E layer and that its intensity is 10^6 a. It is then necessary to clarify the problem of the circuit with the current. If it is assumed that the current is concentrated for the most part in the auroral zone, it can be assumed that the width of the circuit is 300 km, although in actuality the width of the auroral zone varies and apparently changes somewhat at the time of a magnetic storm. With respect to the height of the circuit, if we are speaking of the ordinary E layer, it would be possible to proceed on the basis of the simple layer model, and assume its thickness to be equal to four times the height of the homogeneous atmosphere.

Although a specific sporadic E layer, characterized by a relatively small thickness, is observed at the time of auroras, we must recall what has been said above to the effect that it is entirely possible that there is an inflation of the atmosphere at the time of magnetic storms at the level of the ordinary E layer as well, so that in general we will not be too far from the truth if we assume that the height of the circuit is equal to approximately 40 km. Then the area of the circuit with

the current is about $1.2 \cdot 10^{14}$ cm^2. Expressing the current intensity of the storm in CGSM units of current intensity, we obtain a current density $i = 8.4 \cdot 10^{-10}$ CGSM cm^{-2}.

It is then necessary to find the electrical conductivity in the direction normal to the lines of force of the geomagnetic field. The computations can be made using the following formula (Ref. 5):

$$\sigma_{\perp} = n_e (b_e \sigma_e + b_i \sigma_i), \tag{4.10}$$

where n_e is the electron or positive ion concentration (there are virtually no negative ions at this height), σ_e and σ_i are given by the formulas

$$\sigma_e = \frac{e^2}{m_e v_e}, \quad \sigma_i = \frac{e^2}{m_i v_i}. \tag{4.11}$$

Here e is the elementary charge (in CGSM units); m_e and m_i are the electron and ion masses, respectively; v_e and v_i are the numbers of collisions of electrons and ions with molecules and atoms respectively per second (collisions between electrons and ions do not have appreciable significance at this height), b_e and b_i are found using the formulas

$$b_e = \frac{v_e^2}{v_e^2 + p_{H_e}^2}, \quad b_i = \frac{v_i^2}{v_i^2 + p_{H_i}^2}, \tag{4.12}$$

here P_H is the angular gyromagnetic frequency, obtained using the following expressions:

$$p_{H_e} = \frac{m_e}{Fe}, \quad p_{H_i} = \frac{m_i}{Fe}. \tag{4.13}$$

Here F is geomagnetic field strength. The following values (Ref. 5) can be assumed for a height of about 100 km:

$$\sigma_e = 1.4 \cdot 10^{-18}, \ \sigma_i = 8.5 \cdot 10^{-22}, \ b_e = 1.4 \cdot 10^{-3}, \ b_i = 1.$$

Under normal conditions $n_e = 10^5$, but at the time of strong auroras this concentration can attain 10^6 (Ref. 15). Hence $n_e b_e \sigma_e \simeq 2.0 \cdot 10^{-15}$, $n_i b_i \sigma_i = 8.5 \cdot 10^{-16}$, and $\sigma_\perp = 3 \cdot 10^{-15}$. By substituting the values $\sigma_\perp$ and i into formula (4.9) we obtain $q = 2.3 \cdot 10^{-4}$ $erg \cdot cm^3 \cdot sec^{-1}$. Such thermal energy is comparable to that resulting from elastic collisions when $\boxed{H}$ = 1,000 and v = 1,000 km/sec. This result appears rather probable.

The following mechanisms are responsible for loss of heat from the layer: radiation, molecular heat conductivity, horizontal macroturbulent exchange and convection. We note that the effect of these mechanisms is characteristic not only for gas which is heated by geoactive solar radiation which is excessive in comparison with mean values, but also under ordinary conditions.

Under nonstationary conditions, arising under the influence of solar corpuscular radiation, we can speak of only a certain increase in heat loss associated with an increase of the temperature gradient and corresponding coefficients. The nonstationary state arising as a result of these factors obviously will continue until the additional heat loss is compensated by an additional influx. Then a stationary state again sets in, but with a somewhat higher temperature. The heat loss by radiation already has been taken into account in part in expression (4.8).

A certain part of the Joule heat also will be expended on luminescence in the infrared region, and if it is assumed to be equal to the corresponding part of the energy of primary particles, it will be necessary to discard 60 percent of the value q, since this part of the energy is released in luminescence. Taking the above coefficient into account, the Joule heat expended on heating for the above-mentioned example is $9.4 \cdot 10^5$ $ergs \cdot cm^{-2} \cdot sec^{-1}$; this corresponds to the energy of elastic collisions when there is a concentration of approximately 300 particles at a velocity of 1,000 km/sec or a larger concentration at a lesser velocity.

It is possible that in such an approach to the process of energy loss due to radiation with subsequent luminescence there will be some underestimation of this loss from the total quantity of energy appearing as a result of elastic collisions and an overestimate of the energy loss occurring as Joule heat. However, in principle, the estimate of energy with an allowance for that part that is radiated in the infrared region, made earlier, gives the proper order of magnitude with respect to allowance for heat loss by radiation.

We will now consider the energy loss caused by molecular heat conductivity which in the opinion of certain authors, such as Nicolet (Ref. 16), is the most important loss mechanism. The following energy per 1 cm^3 escapes in 1 sec as a result of molecular heat conductivity

$$\varepsilon_1 = \gamma \frac{\partial^2 T}{\partial z^2}, \tag{4.14}$$

where γ is the coefficient of molecular heat conductivity, z is the vertical coordinate. The horizontal temperature gradient at the level considered is considerably less than the vertical gradient; therefore, the heat loss in a horizontal direction, caused by molecular heat conductivity, need not be considered. The absolute value of expression (4.14) can be represented as

$$\varepsilon_1 = \gamma \frac{\Delta T}{\Delta z^2}. \tag{4.15}$$

The vertical temperature gradient at the level of the E layer of the ionosphere can be assumed to equal 5°/cm (Ref. 17). For the coefficient of molecular heat conductivity, we assume the value $\gamma = 6.2 \cdot 10^4$ $ergs \cdot deg^{-1}$. Then $\epsilon_1 = 3.1 \cdot 10^{-6}$ $erg \cdot cm^{-3} \cdot sec^{-1}$. The coefficient γ is considered constant during the entire time of corpuscular disturbance of the earth's atmosphere. In principle, as already noted, this possibly is not true, but since there is no model of its change at the time of a nonstationary phenomenon, such as corpuscular heating of the atmosphere, such a simplification is inevitable.

Another possible mechanism of heat loss is convection, but it arises under conditions of unstable stratification when the vertical temperature gradient is greater than the adiabatic gradient and temperature decreases with height. At the level considered, however, the temperature increases with height approximately 5° per kilometer, and, therefore, there can be no convection, at least in the first stage of the process. At any rate, it develops only after there is a very considerable heating of a thin layer with total disruption of stratification, characteristic for these levels.

A third possible mechanism of heat loss is horizontal macroturbulent exchange. This mechanism will be effective in a case in which there are large vortices (cyclones and anticyclones) at the level of the lower ionosphere; these vortices are sort of "upper stories" of the corresponding tropospheric formations. Such a situation, as pointed out by L. R. Rakipova (Ref. 18), is entirely probable. It is superfluous to

mention that, as when taking molecular heat conductivity into account it is possible to neglect the heat loss in a horizontal direction in comparison with a vertical direction, when dealing with macroturbulent exchange it is possible to neglect vertical exchange in comparison with horizontal exchange, since the coefficient of the second should be several orders higher than the first. The expression for heat loss due to macroturbulent horizontal exchange has the form

$$\epsilon_2 = 4.19 \cdot 10^7 c_p A \frac{\partial^2 T}{\partial y^2}. \qquad (4.16)$$

Here c_p is heat capacity at a constant pressure, equal under normal conditions to 0.24 $cal \cdot g^{-1} \cdot deg^{-1}$, changing very little in a broad range of temperature and pressure, A is the coefficient of macroturbulent exchange, $\frac{\delta T}{\delta y}$ is the meridional temperature gradient. A simplification has been introduced here: it is assumed that along a circle of latitude in the auroral zone there are no temperature changes under either stationary or nonstationary conditions. It is entirely possible that in the light of the work of A. P. Nikol'skiy (Ref. 19) this assumption is not entirely correct; but it can be accepted as a rough approximation.

For the mechanism of horizontal exchange to operate, it is necessary that there be large-scale vortices at the level considered, i.e., cyclones and anticyclones. According to L. R. Rakipova, cyclonic and anticyclonic disturbances arising in the troposphere can be propagated to the level of the ionosphere in several hours. The dimensions of the vortices do not change appreciably in this process, and have a diameter of about 500 km. Even if, as has been postulated recently, the relationship between the upper and lower layers of the earth's atmosphere is more complex, it nevertheless is entirely probable that tropospheric disturbances "are continued" in some form into the upper layers of the atmosphere.

In order to find ϵ_2 by the use of formula (4.16), it is necessary to determine the coefficient of exchange and the increment of the meridional temperature gradient. The first of these values can be determined using formula (2.6). The mean value of the velocity of the fluctuation $\overline{v}$ can be assumed to be 10^4 $cm \cdot sec^{-1}$, since the wind has approximately the same velocity in the ionosphere as at the height of the E layer (Ref. 19). The velocity of the fluctuations usually should be less than the mean velocity; therefore, the assumed value 10^4 is somewhat too high.

It is obvious that in principle the entire problem should be considered with a variable exchange coefficient. But since the law of its change with time remains unknown, from the very beginning we will proceed on the basis of the maximum value A. This corresponds to formulation of the problem of the theoretical possibility of heating, which we hold here. In addition, the somewhat excessive value of the coefficient of horizontal exchange can to a certain extent compensate the possible failure to take into account heat losses by radiation, and especially compensate the discarding of the component of loss in a direction perpendicular to the meridian.

The assumption is made that the mean value of the mixing path is equal to 10^8 cm, which somewhat exceeds the mean diameter of the vortex. However, considering the distance between vortices (the so-called "origin path", Ref. 20), it can be assumed that 10^8 is close to the true value. Since at the height of the E layer density is equal to $2 \cdot 10^{-10}$ $g \cdot cm^{-3}$, then $A = 2 \cdot 10^2$ $g \cdot cm^{-1} \cdot sec^{-1}$. It is necessary to take the turbulence spectrum into account, at least approximately.

In addition to vortices of such large scale as cyclones and anticyclones at the level of the E layer, there can be smaller vortices with a diameter of several tens of kilometers or less. The increment to the value A caused by these smaller vortices cannot be very large, since it decreases with a decrease of characteristic dimensions. It was assumed to be a component of 50 percent of A, but it is entirely possible that this is too high; but here we can cite the same arguments used in selecting the velocity of the fluctuation. An effort will be made to obtain the least favorable conditions for an increase of temperature as conditions for the theoretical possibility of some heating in general.

Assuming the half-width of the auroral zone to be 150 km and the temperature drop, assuming a standard distribution, to be 1° between the middle of the zone and its northern periphery, it is possible to find the meridional temperature gradient, and therefore the heat loss from the zone due to macroturbulent horizontal exchange. The computations give $1.34 \cdot 10^{-5}$ $erg \cdot cm^{-3} \cdot sec^{-1}$, which almost exceeds by a factor of 4 the loss due to molecular heat conductivity. In the light of Nicolet's concept that the latter is the most effective mechanism (Ref. 16), such a result can indicate that the coefficient of macroturbulent exchange has an excessive value. However, Johnson (Ref. 21) also has found that the horizontal loss of heat is associated not with exchange, but with advection, and is a more effective mechanism of heat transfer than molecular heat conductivity in a vertical direction.

The formulated problem can be reduced to the numerical solution of the nonstationary heat balance equation. It is necessary to find: (1) the increment of temperature leading to the establishment of a stationary state on a higher temperature background; and, (2) the time within which this new state is established. If the heat increment results from elastic collisions of protons with air particles, the nonstationary heat balance equation is written as follows:

$$\frac{dq}{dt}=0.4\,\frac{m\,[H]\,v^3}{2S}-c_pA\,\frac{\partial^2 T}{\partial y^2}-\gamma\,\frac{\partial^2 T}{\partial z^2}, \tag{4.17}$$

or, if we convert to finite increments:

$$\frac{\Delta q}{\Delta t}=0.2\,\frac{m\,[H]\,v^3}{S}-c_pA\,\frac{\Delta T}{\Delta y^2}-\gamma\,\frac{\Delta T}{\Delta S^2}, \tag{4.18}$$

where S is the height of the homogeneous atmosphere.

Since the expansion of the zone in a horizontal plane can be neglected, $\frac{1}{\Delta y^2}$ = const. The value $\frac{1}{\Delta s^2}$ is obtained for the initial conditions. We will select a certain time interval, such as 10 seconds, and assume that in the course of the first such interval there is no loss associated with an increase of gradients. Then the entire energy increment is expended on an increase of temperature. By increasing the first term of the right-hand side of equation (4.18) by a factor of 10, that is, taking the heat influx for 10 seconds, it is possible to determine the temperature increment from the relation

$$\Delta T=\frac{2}{3}\,\frac{\Delta E}{Nk}, \tag{4.19}$$

where ΔE is the influx of energy during the course of the first interval; N is the number of molecules in a unit volume; and, k is the Boltzmann constant.

The temperature increase has a double result. First, it is reflected in the value S, and therefore ΔS, since it is assumed that ΔS = S, and second, in the value ΔT. In order to compute the increase of temperature during a time corresponding to the second interval, that is, during the second 10 seconds, it is necessary to compute a new value of the first term of the right-hand side, and also of its remaining terms, proceeding from the value ΔT computed using formula (4.19). Determining Δq in this way for 1 second during the second interval by this method, we increase the derived value by a factor of 10 and obtain a new value, ΔT, by use of formula (4.19).

The temperature increment in 20 seconds, that is, the deviation of temperature from an initial equilibrium state, will be equal to the sum of the temperature increments during the first and second 10 seconds. We then use a third interval, in which ΔT is taken equal to the sum of these values during the first and second 10 seconds. Finding the new value Δq for 1 second, and then for 10 seconds, in the course of the third interval we determine the increment of temperature during the third 10 seconds, and adding this to the already determined temperature deviation from the initial equilibrium state, we perform the next step, etc.

The establishment of a new equilibrium state will be characterized by the fact that the temperature increment tends to zero and that its general increase ceases. For practical purposes the approximations can be ended when the temperature increase drops to 0.1 of the value it had during the first 10 seconds.

The computations were made using the following values: particle concentrations 1, 10, 100 and 1,000 cm^{-3}; particle velocities: 1,000, 2,000 and 3,000 km/sec; proton mass $1.66 \cdot 10^{-24}$ g; height of the homogeneous atmosphere $S = 10^6$ cm (initial value); specific heat capacity $c_p = 0.24$ cal/deg $= 4.185 \cdot 0.24 \cdot 10^7 = 10^7$ ergs/deg; coefficient of horizontal macroturbulent exchange $A = 3 \cdot 10^2$ $g \cdot cm^{-1} sec^{-1}$; coefficient of molecular heat conductivity $\gamma = 6.2 \cdot 10^4$ $g \cdot cm^{-1} sec^{-1}$; half-width of the auroral zone $\Delta y = 1.5 \cdot 10^7$ cm; number of air molecules per 1 cm^3 at the level of the E layer of the ionosphere $5 \cdot 10^{12}$; Boltzmann constant $k = 1.37 \cdot 10^{-16}$ $erg \cdot deg^{-1}$ $mole^{-1}$; $1/\Delta y^2 = 4.45 \cdot 10^{-15}$ cm^2; and, $1/\Delta z^2 = 10^{-12} cm^{-2}$.

In computation of the height of the homogeneous atmosphere, it was assumed that $\mu = 25$; $g = 950.6$ cm^{-2} (height 100 km).

The results of the computations are given in Table 29. The numerators of the fractions in this table show the temperature increment in degrees, responsible for establishment of the new stationary state. The denominators express the number of seconds necessary for the establishment of this stationary state. As shown by Table 29, at a more considerable velocity of the particles the stationary state is established somewhat more rapidly, but it is obvious, at a higher level.

Table 29

Quantity of protons, cm^3	Velocity of protons, km/sec		
	1000	2000	3000
1	Virtually no heating		
10	$\frac{0.2}{170}$	$\frac{1.5}{120}$	$\frac{6}{80}$
100	$\frac{2.5}{190}$	$\frac{17}{150}$	$\frac{50}{120}$
1000	$\frac{20}{80}$	$\frac{130}{140}$	$\frac{300}{100}$

With respect to the temperature increment, when there is a concentration of several thousand particles having a velocity of about 1,000 km/sec, the temperature increment in the E region of the ionosphere is several tens of degrees. If we assume that Joule heat is the mechanism responsible for heating, we can return to the already cited example.

When the current intensity of the storm is 10^6 a, when the circuit is the cross section of the auroral zone and conductivity across the lines of force is $3 \cdot 10^{-15}$, the influx of heat, as already pointed out, will be $2.34 \cdot 10^{-4}$ erg$\cdot$cm^{-3}sec^{-1}. This corresponds to the heat influx from elastic collisions when there is a concentration of 1,000 particles$\cdot$cm^{-3}, and a velocity of 1,000 km/sec. The temperature increment in this case will be about 20°.

It should be noted that in actuality the system of currents of a magnetic storm to one degree or another involves the entire earth, and is not concentrated solely in the auroral zone. The system of currents involved in the above computations should be considered as a first approximation.

Taking into account the assumptions made in the development of the above-mentioned model, it should be assumed that the heatings cited in Table 29 are the lower limits. Recently published investigations by Cole, on the other hand, give the upper limit of heating. Cole concludes that at the time of magnetic storms and auroras there is an inflation of the atmosphere, leading to an increase of the height of the homogeneous atmosphere by a factor of 5-10, and that such an order of increase of the height of the homogeneous atmosphere should occur at heights of about 150 km (Refs. 22, 23).

The heat flux arriving from the auroral zone, according to Cole, should increase by a factor of 10 at the time of strong magnetic storms. In addition to Cole's work, there is an interesting recently published communication by King-Hele (Ref. 24), who investigated the change in height of the homogeneous atmosphere at heights of 200-450 km on the basis of data on artificial earth satellite braking. He obtained data for the following epochs: (1) beginning of 1958; (2) beginning of 1959; and, (3) 1960-1961. Table 30 from the study of King-Hele gives data on the height H of the homogeneous atmosphere, molecular weight at the corresponding heights on the basis of data published by Kallman-Bijl (Ref. 25), and on temperature. As shown by the table, at the same height and coinciding closely with the level of the F layer, the temperature from 1958 through 1960-1961 fell by approximately 500°, which exceeds by a factor of approximately 2 its decrease on the basis of ionospheric data (see above, Figure 38). By comparing Figure 38 with the data in Table 30 we can see that the temperature value as indicated by artificial earth satellite braking data exceeds by a factor of approximately 2 the corresponding value determined on the basis of ionospheric data, which is not a new discovery (Ref. 26).

For completeness we will cite still another recently published investigation by Paetzold (Ref. 27), concerning the relationship between the heating of the upper atmosphere at heights greater than 400 km, and solar active radiation. This investigator was able to separate the heat flux caused by wave and corpuscular solar radiation, and to establish the relationship between each of the thermal increments and solar radio emission at a wavelength of 10.7 cm.

Table 30

	Beginning of 1958			Beginning of 1959			During 1960-1961		
	height, km								
	200	250	300	250	350	450	270	350	450
H, km	51	61	66	50	61	64	45	53	57
. . . .	27	26	24	26	22	19	25	22	19
T^{o} K	1530	1720	1680	1400	1400	1230	1220	1220	1230

CHAPTER 5
MANIFESTATIONS OF SOLAR ACTIVITY IN THE LOWER LAYERS OF THE EARTH'S ATMOSPHERE

Not only in astronomy and geophysics, but indeed in the fields studied by the other sciences, there are few such problems about which so much has been written as problems involved in the relationship between solar activity and weather and climatic phenomena. We already have mentioned in the foreword that in this sphere, where the problems of astrophysics, aeronomy, meteorology, climatology and hydrology are intertwined closely, a correct and at the same time sufficiently new approach to the problem is impossible without a thorough and critical analysis of what has been done earlier in this field.

The sun-troposphere problem is a name which can be assigned in brief, and to a certain degree, arbitrary form to that field of investigation discussed in this chapter, and because of its extensive nature it should be somehow subdivided or classified. One of the criteria for making such a classification is the assignment of large groups of investigations to specific historical stages. It is possible to define several such stages in the development of the sun-troposphere problem. At the later stages it is possible to separate the comparison of solar activity with one meteorological element from comparisons with others, right up to comparisons of solar activity with the characteristics of general circulation of the atmosphere, indicative of the high level of development of solar-meteorological investigations.

Finally, an important criterion in the classification of investigations of the sun-troposphere problem is the interval of time within which the comparison of solar and meteorological (hydrological) phenomena is carried out. From this point of view investigations of the sun-troposphere problem can be divided into the following categories.

1. Comparisons within intervals of time exceeding the 11-year and 22-year cycles of solar activity. This category includes determinations of variations in climate, the discharge of rivers, the level of reservoirs, etc., the 80-90-year cycle, and also the longer cycles. This category of comparisons is characteristic in that in essence such investigations often result in the discovery of new solar cycles. Variations of climate and the hydrological regime with a duration of several centuries or even thousands of years can be caused by long cycles of solar activity.

2. Comparisons within the 11-year and 22-year cycles. Here we must consider where it is rational to put comparisons within the so-called Brükner cycles. We feel that it is desirable to assign such investigations to class 1, assuming the corresponding cycles to be either parts of the 80-90-year cycle, such as a 44-year cycle, or to cycles of the same order, since the 80-90-year cycle sometimes is of appreciably shorter duration (see Chapter 1).

3. Twenty-seven-day cycles in meteorological phenomena.

4. "Individual" comparisons of solar activity with meteorological phenomena. However, we are not referring to truly individual comparisons, that is, thorough determinations of the heliophysical and geophysical conditions at the time of one or more outstanding meteorological phenomena, but instead statistical investigations of a large number of similar meteorological phenomena compared with some particular index of solar activity, or with some geophysical index whose relationship to solar activity is beyond question (such as with geomagnetic data). Individual comparisons, in the true sense of the word, have been made far more rarely.

A special aspect of the sun-troposphere problem is the group of problems associated with the mechanisms by which solar activity influences meteorological phenomena.

In our earlier reviews of investigations of this problem (Refs. 1, 2) the only basis used in the classification of solar-meteorological comparisons was separation by the principle of those intervals of time within which the comparison was made. At the present time, since the number of investigations of the sun-troposphere problem has become extremely great, it appears inadequate to make a classification using this single criterion only.

It is now necessary to define specific stages in development of the problem. Approximately the following historical groupings can be defined: (1) from the discovery of sunspots to the discovery of the solar cycle; (2) from the discovery of the solar cycle to the appearance of the first "individual" comparisons of solar and meteorological phenomena; (3) from the epoch when investigations devoted to manifestation of long-term cycles in climate and "individual" comparisons began to the present time; and, (4) current investigations.

We can mention the principal studies that should be regarded as the beginning of the corresponding new periods of sun-troposphere comparisons. In the first stage we have a beginning with the letter from B. Balliani to Galileo; in the second, the basic study made by Köppen (Ref. 3); in the third, the series of studies made by Soviet researchers (M. S. Zhukov, P. P. Predtechenskiy, and others); and in the fourth, the

investigation of the Duells and recent investigations of the mechanisms governing the relationship between solar activity and the troposphere (A. A. Dmitriyev and L. R. Rakipova in the USSR, Spar in the United States, etc.).

It is in accordance with this classification that we will review the research which has been done in this field.

Section 1. First Stage in the Development of Sun-Troposphere Investigations

The specific approach of natural philosophy characteristic of the times is representative of the first stage in the study of the relationship between apparent manifestations of solar activity and meteorological processes. As already mentioned, immediately after the discovery of sunspots by optical instruments, Battiste Balliani (Ref. 4) wrote Galileo that these spots can be regarded as "coolers", and when there is a large number on the surface of the sun it is possible to expect lower temperatures on the earth. Riccioli (Ref. 4) compared the temperature of several places in Italy in July-September, 1632 and 1642.

In the first case, no sunspots were visible during these months (this was an epoch two years prior to the regular minimum of the 11-year cycle, which, however, was unknown until considerably later). The weather conditions in July-September, 1632, were characterized by a severe drought. In the second case, that is, in 1642, the number of visible spots was very great (this was three years before the next minimum—apparently of a higher 11-year cycle). In this case the onset of summer was characterized by low temperatures.

Deschale (Ref. 4), applying the relationships noted by the Italian investigators to Paris, pointed out in 1690 that the conclusions drawn by Balliani and Riccioli are not of a universal character, that is, there can be geographic areas in which an increase in solar activity does not lead to cooling, but to aridity, and vice versa, a decrease in solar activity causes an increase in precipitation. Chr. Wolf (Ref. 5), who investigated weather conditions in 1709 (which fell in an epoch of increasing solar activity after a long period of low activity late in the 17th century and at the very beginning of the 18th century), when not a single spot was observed in the solar northern hemisphere over a period of more than 30 years, concluded that the total area that sunspots can occupy is too insignificant to influence solar radiation.

Thus, there existed certain concepts concerning solar-meteorological relationships even before the well-known investigation of W. Herschel with which it usually is assumed that the investigation of the sun-troposphere problem began. Herschel (Ref. 6) discovered a relationship

between the prices of certain agricultural commodities and sunspot number. Essentially he concluded that years with many sunspots are at the same time warmer and more favorable for agriculture.

The next investigation worthy of mention is Arago's study of the dependence of the temperature of Paris on sunspot number (Ref. 7). He succeeded in noting that in years with a large number of sunspots, the temperature averaged 0.3^{o} lower than in years in which there were few spots. The mean annual amplitude of temperature was greater when there were few spots than when there were many spots, and was less in years in which there were few spots. At the same time, years with a large number of spots are years with abundant precipitation. These results were confirmed by Gautier (Ref. 8), who had at his disposal temperature observation series for a number of stations.

Gautier found that the temperature at Paris in years with a large number of sunspots was 0.6^{o} lower and at Geneva 0.3^{o} lower than in those years in which there were few spots. It should be noted that in his investigation Gautier studied the same series of years (1820-1843), the study of which enabled Schwabe to discover the solar cycle. In this respect it would be more correct to assign Gautier's work to investigations in the second period of development of the sun-troposphere problem.

If we summarize briefly the results of the first period of study of solar-tropospheric relationships, we can note the following: even in that period one of the distinguishing characteristics of this aspect of the sun-earth problem was becoming clear, that is, the dependence of the result on the geographic area within which the comparison is made. It is this factor that explains the apparent contradiction between the results obtained by Herschel and Gautier, as was pointed out by R. Wolf (Ref. 9). In actuality, the conclusion that years with large numbers of sunspots are favorable for agriculture was not made by Herschel on the basis of data for Paris or Geneva, and, therefore, already in this difference in the mentioned authors' results we can see how dependence on geographic location is manifested.

With respect to the first period of development of investigations of the sun-troposphere problem, it is impossible to pass over the interesting conclusion drawn by Chr. Wolf on the basis of the concepts of natural philosophy that sunspots, due to their small dimensions and, in general, their rather limited number at any particular time, cannot influence solar luminosity (in the sense of visible radiation). As is well known, this conclusion is confirmed fully by later investigations.

Section 2. Second Stage in Development of Solar-Tropospheric Investigations

As already mentioned, the most typical of the studies that initiated the second period of study of solar-tropospheric relationships can be considered the investigation of Köppen (Ref. 3), published in 1873, dealing with the variation of temperature in the 11-year cycle. Historically, the Köppen study was preceded, however, by a number of investigations which merit mention.

We should first mention certain comparisons made by R. Wolf himself, who in 1852 analyzed Zurich data for the period 1000-1800 (Ref. 9). Wolf concluded, in contradiction to the results of Arago and Gautier, but in agreement with Herschel, that rainy and stormy weather was characteristic of years of minimum solar activity, and dry and warm weather (but at the same time not too arid, that is, generally favorable for agriculture) was characteristic of years of maximum solar activity. In order to clarify the noted contradiction, Wolf in 1859 used data on temperature at Berlin, dividing into two parts the series of observations from 1760 through 1847 that he had at his disposal. The breakdown was from 1760 through 1802, and from 1803 through 1847. It was found that Herschel's conclusion is valid for the first period (to which, incidentally, part of Herschel's data applies), but Gautier's conclusion is correct for the second period.

Currently, when it is possible to conceive of the manifestation of an 80-90-year cycle in climate, such a difference becomes understandable. The first of the epochs considered by Wolf includes a maximum and descending branch of an 80-90-year cycle, and the second of the epochs includes its minimum and the ascending branch. Köppen's results are shown in Figure 39. The upper curve represents the inverted curve of mean annual Wolf numbers; the lower curve shows the variation of mean annual temperatures of the tropics. The figure shows that the amplitude of the variations of mean annual temperatures is close to 0.5°; the curves of temperature variation and the inverted curve of variation of Wolf numbers are parallel.

Figure 39 also shows data for a later epoch than on the original Koppen graph. This extension of Köppen's diagram was made by Nansen and Helland-Hansen (Ref. 10); their principal objective was clarification of the behavior of the temperatures of the temperate latitudes in the 11-year cycle. The corresponding curve is shown in the lower part of the figure. In this case there is no such well-defined relationship as can be seen in the case of tropical temperatures. The upper curves indicate a certain displacement relative to one another; the maxima and minima of tropical temperatures set in after the minima and maxima of solar activity respectively. In this connection Köppen notes: "Just as the greatest cold sets in after the winter solstice, and the greatest heat after

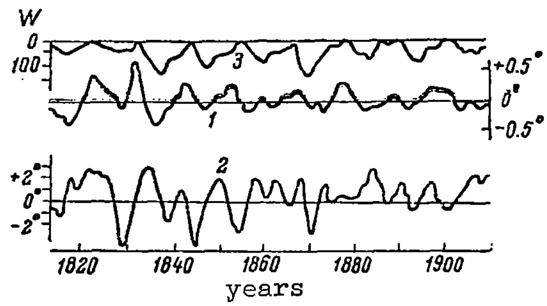


Figure 39. Air temperature and solar activity. 1, mean annual temperatures of the tropics; 2, mean annual temperatures of the temperate latitudes; 3, Wolf numbers

the summer solstice, it can be assumed that something similar happens in relation to the manifestation of solar activity in tropical temperatures" (Ref. 3, p. 265).

Many studies similar to that of Köppen's were made during the period 1880-1930, and it is impossible to discuss each of them even briefly. For the most part the studies were of the manifestation of the 11-year cycle, with fewer studies of the 22-year cycle.

Following purely graphic comparisons of the variation of solar activity (on the basis of Wolf numbers) with some particular meteorological index for a particular region, investigations appeared in which harmonic, periodogram and especially correlation methods were employed. The most significant results in this period of study of solar-tropospheric relationships were as follows.

1. A dependence was established between a whole series of weather elements and the phase of the 11-year solar cycle. However, the amplitudes of the variations of these elements were small.

2. Clayton compiled maps of the distribution of centers of high and low temperature over the earth's surface, as well as similar maps of increase and decrease of pressure and increase or decrease of the quantity of precipitation in the 11-year cycle (Ref. 11). He compared the mean values of the corresponding elements in an epoch of the maximum (year of the maximum ± 2 years), and in an epoch of the minimum (year of the minimum ± 2 years), in general, for five 11-year cycles (Nos. 11-15). On these maps he found for the most part a somewhat regular, although complex distribution of regions of positive and negative correlation with solar activity.

3. Walker (Ref. 12) and Brooks (Ref. 13) computed a large number of correlation coefficients of Wolf numbers with various weather elements. The first of these authors established, in particular, that low pressure in the equatorial zone becomes particularly low in epochs of the maxima of the cycle. At the same time, subtropical anticyclones become strongest. Brooks established a close relationship between precipitation in certain drainage areas of the equatorial zone (such as southeastern Africa) and solar activity. For example, the well-known 11-year variation in the level of Lake Victoria was established. In addition, it was Brooks who established a close relationship between the number of thunderstorms over the entire earth and Wolf numbers. The regional relationships between these phenomena and solar activity in some places were extremely close (Ref. 14). For example, the correlation coefficient in Siberia for the years 1888-1924 was +0.89.

4. Ye. Ye. Fedorov (Ref. 15), V. Yu. Vize (Ref. 16) and Walker (Ref. 12) discovered the so-called law of accentuation of pressure fields, which provides that with an intensification of solar activity in the 11-year cycle, there is an intensification of the pressure field of a particular sign; in other words, in regions occupied by low-pressure areas there is a further decrease in pressure, that is, a deepening of the cyclones, whereas in anticyclonic regions there is a further increase of pressure, that is, the anticyclones are intensified even more. The conclusions drawn by Walker, mentioned above, are also illustrated in the example of the equatorial zone of low pressure and subtropical anticyclones, confirming the general law of accentuation of pressure fields.

5. It was this same stage of the study of solar-tropospheric relationships that saw the appearance of the first noteworthy hypotheses of the physical nature of these relationships. Particularly important in this respect was the hypothesis of variations of the solar constant, whose authors were representatives of the so-called Smithsonian school (United States): Abbot, Fowle, Clayton and others (Ref. 17). These investigators assumed that variations of the solar constant are completely real and associated closely with solar activity. However, variations of the solar constant should be related to meteorological phenomena. Nevertheless, even at this stage of development of the sun-troposphere problem, serious doubts were expressed concerning the reality of such a mechanism of solar-meteorological relationships. The principal argument of opponents of the concept of the Smithsonian school was that with improvement of methods of determining the solar constant, its determined variations would become increasingly smaller. It could therefore be concluded that if such variations in general do exist, they should be very small, possibly falling in the limits of errors in observation. This same period saw the appearance of another hypothesis on the mechanism whereby solar activity influences meteorological phenomena. This was the so-called Humphreys hypothesis of an ozone shield (Ref. 18).

Humphreys felt that the solar short-wave radiation, intensifying toward the maximum of the 11-year solar cycle, leads to an increase of the ozone content in the earth's atmosphere. At the same time there is an increase in the percentage of the earth's infrared radiation that is held back by the atmosphere, and results in its warming. However, later investigations revealed that such a model of the influence of solar activity on the troposphere cannot withstand criticism.

6. The first more or less satisfactorily executed "individual" comparisons of solar and geophysical phenomena appeared at the end of the considered period of investigations on the sun-troposphere problem.

Section 3. Third Stage in the Development of Investigations on the Sun-Troposphere Problem

This stage, beginning at the threshold of the 1930's, is characterized by the development of investigations of the 27-day cycle in meteorological phenomena, and especially the development of investigations of the type of individual comparisons of solar and meteorological phenomena. As already mentioned, it is incorrect to refer to these comparisons as "individual". These constitute the same statistical investigations, but they cover a fewer number of years than those in which the objective is to study manifestations of the 11-year cycle in meteorological processes.

In most cases, the unit of reckoning in investigations of this type is the day. At the same time, in this third stage of the study of sun-earth relationships investigators did not forget to study the manifestations of the 11-year and other solar cycles in meteorological processes. We can note here that it was established that in a considerable number of regions the principal solar cycle manifested in the lower layers of the earth's atmosphere was the 22- and 80-90-year solar cycles, not the 11-year cycle.

In a number of respects the third stage of development of solar-tropospheric investigations passes directly into the fourth, or recent stage, whose principal characteristic is that the greatest attention is devoted to problems of the mechanisms involved in the relationship between solar and meteorological phenomena. The third stage, whose termination can be set arbitrarily at 1950, is characterized by the development of investigations of the manifestation of the cycle of solar rotation in the lower layers of the earth's atmosphere, and the development of the statistical aspect of ideas serving as the basis of the work of Memery (Ref. 19), Aufsess (Ref. 20) and others concerning the individual effects of active solar centers in tropospheric phenomena.

This point is explained with the following example: Memery uses a simple catalog of cases in which the passage of a spot group across the central meridian, or its emergence at the western limb or appearance at the eastern limb of the disk (the possible geophysical effects induced by the presence of active solar formations at the limbs will be discussed below), or, finally, rapid changes in the dimensions or configuration of these groups lead to changes in macroweather. Later investigators considered similar problems, but statistically, applying statistical methods, such as the superposing of epochs, to the corresponding data.

A wide variety of investigations was undertaken by the repeatedly mentioned group of Tashkent geophysicists (M. S. Zhukov, P. P. Predtechenskiy and K. V. Brodovitskiy (Ref. 21). The most remarkable aspect of their work is a study of the manifestation of the 27-day period of solar rotation in a number of weather elements and the hydrological regime, and also in the frequency of recurrence of tropospheric macroprocesses. The investigation method involves the construction of Carrington calendars of a particular macrometeorological phenomenon and a certain solar or geophysical index, the latter in a case in which there is no doubt as to its relationship to solar activity (such as geomagnetic activity).

By summing the indices in the Carrington 27.3-day calendar for a certain interval of time, Zhukov and Predtechenskiy in a number of cases obtained graphic pictures of the relationship between the two. An example of the application of this method is the determination of the sums for the days of the 27.3-day calendar of the number of cases of appearance of polar maxima in the Spitzbergen and Novaya Zemlya zone for the years 1930-1933 in comparison with the similar sums of the number of days with magnetic storms for these same years. Both the first and second curves reveal within the limits of the 27.3-day cycle two gently sloping peaks falling on the 4th-10th and 17th-24th days of the calendar, respectively.

The coincidence of the days of the solar calendar, that is, of solar longitudes, in the Carrington system responsible for the two phenomena can be seen on the unsmoothed curve, but especially on the smoothed curve. There is an equally clear relationship between solar activity and the number of days with polar maxima in all sectors of Eurasia and the Spitzbergen-Novaya Zemlya region. Solar activity is represented here by the intensity of calcium flocculas, in common use at one time. In the study of the Tashkent geophysicists, these data are cited for the entire No. 16 11-year cycle. According to the data for meteorological stations of Central Asia for 1926, precipitation and thunderstorms have maxima and minima within the 27.3-day calendar which coincide satisfactorily with the corresponding maxima and minima of Wolf numbers in the central zone of the solar disk and the international indices of geomagnetic activity.

In certain cases the curve of the solar index in the 27.3-day calendar was displaced relative to the meteorological index. This occurred when the phenomenon characterized by the latter lags relative to the most favorable position of an active region of the sun in the sense of its effect on the earth's atmosphere. An example of this kind is the 27.3-day calendars of calcium flocculas in the central zone of the sun, on the one hand, and atmospheric pressure at Tashkent in 1936 on the other. The displacement of one curve relative to the other in this case is 4 or 5 days, and the maxima of atmospheric pressure at Tashkent lag in relation to the maximum development of calcium flocculas. This situation should cause no surprise if it is remembered that an increase of pressure in Uzbekistan often is caused by Arctic intrusions, and Arctic air requires several days to travel from the shores of the polar seas to the neighborhood of Tashkent.

The expression within the limits of the Carrington cycle of the two large maxima of frequency of recurrence of intrusions of cold air masses into Central Asia in 1934-1936 is rather clear. Equally well expressed is the parallel course of the 27.3-day curves of the number of polar maxima for 1913-1933 and Wolf numbers for 1917-1935. A less satisfactory pattern is obtained for the variation of atmospheric pressure at Tashkent during the period 1923-1935, and especially for the total precipitation for this station for these same years. At the same time certain hydrological indices give a very clear 27.3-day variation and a clearly expressed relationship with solar activity.

These indices include the discharges of the Syr-Dar'ya River, associated with the arrival of water from the glaciers of the Tien Shan and from an extensive drainage basin. Such "accumulative" indices often are the most representative when determining the relationship between a hydrometeorological phenomenon and the solar activity which it describes. In particular, this explains the success of the method of construction of integral difference curves used recently for study of solar-meteorological relationships.

The melting of glaciers and precipitation in certain regions can reflect an intensified interzonal turbulent heat exchange, which, as will be shown below, is associated with solar activity. Each more or less prolonged and strong increase of macroturbulent heat exchange, accompanying the transport of warm air masses, leads to an intensification of the melting of glaciers and to an increase in river discharge.

Completely satisfactory results have been obtained by Tashkent geophysicists in the processing of a 27.3-day calendar of dates of a catalog of anticyclonic storms of the Caspian Sea (Ref. 22). The maxima of the frequency of recurrence in this case are displaced somewhat relative to the days of the solar calendar, depending on the direction of arrival of the anticyclone causing the Caspian storm.

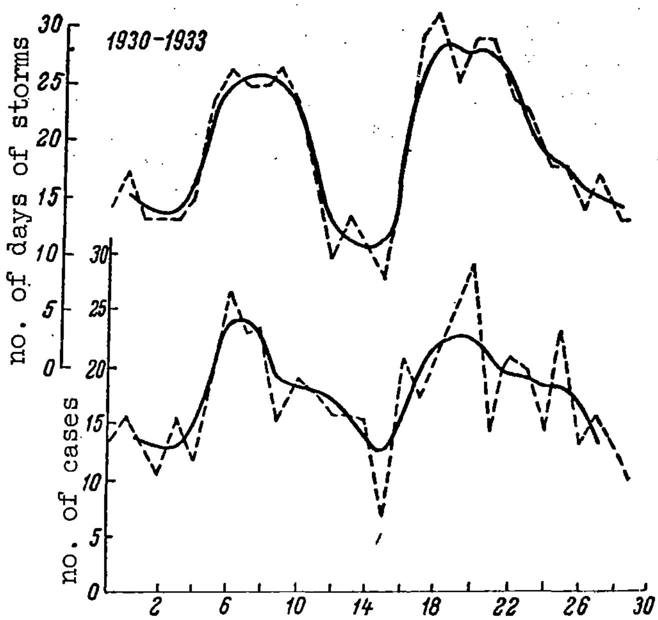


Figure 40. Polar maxima and magnetic storms for 1930-1933 in the 27-day cycle (according to P. P. Predtechenskiy and K. V. Brodovitskiy)

As an illustration we will show several curves obtained by M. S. Zhukov, P. P. Predtechenskiy and K. V. Brodovitskiy. Figure 40 shows the sums of the numbers of polar maxima in the neighborhood of Spitzbergen and Novaya Zemlya (lower curve) and magnetic storms (upper curve) according to the 27.3-day calendar for 1930-1933. Figure 41 shows the 27.3-day curves of the sums of indices of calcium flocculas for 13 solar rotations in 1936, and the sums of negative temperatures over Pavlovsk as indicated by aerological sounding to heights of 5,000 m. In this case we can see some displacement of one of the curves relative to the other; the nature of this displacement was determined more precisely at a later date in one of the author's studies (Ref. 23).

A thorough investigation of the 27.3-day cycle in meteorological phenomena represents an appreciable innovation, and should be considered a major contribution of the Tashkent investigators. They also drew certain conclusions concerning the manifestation of the 11-year cycle in meteorological processes, primarily in Central Asia. These will not be discussed here. While giving due credit to the studies of M. S. Zhukov and P. P. Predtechenskiy it is impossible to pass over certain deficiencies in their investigations.

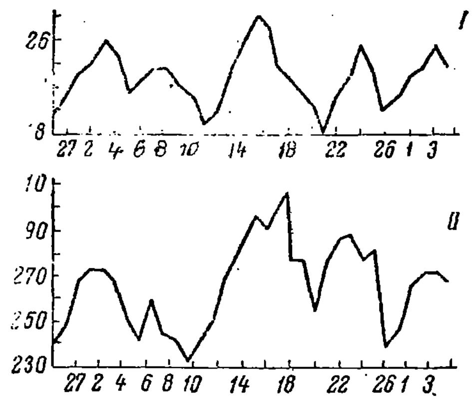


Figure 41. Dependence of temperature of the free atmosphere on solar activity according to P. P. Predtechenskiy and K. V. Brodovitskiy. I, calcium flocculas (international indices); II, temperature at Pavlovsk at height of 5,000 m

We note first that there is an absence of statistical criteria of the reality of the results. It is true that the problem of evaluation of the results obtained using solar calendars still has not been solved in its entirety. There apparently are three different approaches to its solution. The first approach involves the evaluation of the summary curve, representing the sums of the frequencies of the particular index for the days of the 27.3-day calendar or the curve of their mean values, which differs formally from the summary curve only in scale and the number of its extrema. One of the simplest criteria for the solution of this problem has been proposed by B. P. Veynberg (Ref. 24). In essence, when there is a particular number of observations, that is, a particular number of classes of the argument, the number of extrema (maxima and minima) has a definite theoretical error. A table compiled by Veynberg gives the value of this error for a number of observations from 5 to 20 and for a theoretical number of extrema from 3 to 13; in this range the error varies from 0.492 to 1.24. If the number of terms of the series is greater than 20, as occurs in the case of the solar calendar, its random character is investigated by parts; each of these parts should include less than 20 terms. The difference between the observed and theoretical number of extrema is then determined, and if it is not less than three times greater than the theoretical probable error, the series can be considered nonrandom.

It should be noted that in studies made later than those at Tashkent, and in which the solar calendar method was used, the authors use three-day smoothing, and as a result the number of classes of the argument was decreased by a factor of 3, that is, in place of 27 days there were 9 three-day periods, and the frequency of the phenomenon in each of them was obtained as the sum of these frequencies for three successive days of the solar calendar.

Such a decrease in the number of classes of the argument leads to a decrease in the theoretical number of the extrema. One of the shortcomings of the Veynberg criterion is that it is impossible to estimate the confidence coefficient. The Gleissberg criterion described in Ref. 25, similar to the foregoing, does not have this shortcoming. The probability that the observed variation has a cyclic character is determined by the Gleissberg method from a special table that includes not only the number of terms of the series but also the number of extrema, in the same sense as in the Veynberg criterion. The mathematical bases of the Gleissberg criterion are related to the solution of one of the problems in combinational analysis. They have been subjected to doubt by certain mathematicians. Nevertheless, the Gleissberg criterion, like Veynberg's, has the shortcoming that the size of the statistical population is not taken into account in explicit form. An essentially different method was proposed by the author in 1949 (Ref. 26).

We have intentionally discussed in some detail the problem of estimating the confidence coefficient of the results obtained using solar calendars, not only because of the importance of this problem in itself, but also in order to show that at the time when the Tashkent geophysicists made their investigations, the problem of the statistical checking of the results of this type of work still was in ints initial stage of development.

One of the ideas considered in the studies of M. S. Zhukov and P. P. Predtechenskiy is the concept of rhythms of state and rhythms of position. Included in the first are the 11-year and 22-year cycles. The Tashkent geophysicists still have not studied cycles of great duration. Included in the rhythms of position are the semiannual rhythm caused by change of the heliographic latitude of the projection of the earth on the sun, which latitude, as is well known, attains its maxima in March and September and its minima in June and December.

Also included in such rhythms is the 27.3-day rhythm, which has been subjected to investigation by Tashkent geophysicists. It should be noted that they also have detected a 13-14-day rhythm, two of which make up one 27.3-day rhythm. The nature of the 13-14-day rhythm is related to the fact that the principal clusters of active centers on the sun reveal a tendency to concentration near diametrically opposite Carrington longitudes. In geophysical and particularly in tropospheric

phenomena, this tendency is manifested still more clearly than on the sun itself. In certain cases the rhythm can be of nine days duration. This will occur when there are three active centers on the sun, situated 120° apart in longitude.

It should be noted that the Tashkent investigators operated on the assumption of a linear and quite rapid propagation of geoactive radiation from the sun to the earth. It was assumed that the radiation is propagated for the most part in the form of narrow streams which can be incident on the earth only when there is sufficiently direct orientation in the direction of the latter. Successful conditions for such incidence should occur when an active center at the time of an "explosion", that is, the release of activity by the center, is near the center of the visible solar disk. Such a concept was very widespread in the 1930's.

The Tashkent geophysicists assumed that if synoptic processes in Central Asia in certain cases can lag several days behind the date of passage of an active center across the solar central meridian, this can be attributed solely to the distance of Central Asia from those geographic regions in which the effect of solar activity is manifested directly. The travel time of corpuscular radiation from the sun was assumed to average 26 hours. However, it cannot be assumed unconditionally that only the solar central meridian is of decisive importance in the problem of solar-terrestrial, and particularly in solar-tropospheric, relationships. It is possible that in certain cases in which corpuscular streams are propagated slowly, the most favorable position of an active center on the solar disk will not be on the central meridian. In the case of slowly moving particles it is also necessary to take into account solar rotation and its gravitational effects on the corpuscular stream (Refs. 27, 28), which considerably increases the time of its propagation in interplanetary space.

It is not impossible that specific places on the solar disk where a center of activity can have the best "orientation" in the direction of the earth are the limbs of the solar disk, as at first glance does not appear paradoxical. This viewpoint was defended in his time by Memery (Ref. 19). He felt that the individual effect of a spot, spot group or facula on meteorological phenomena takes place during the formation of these elements of solar activity or at the time of emergence of a spot group and also a facula field from behind the eastern limb of the disk and at the time of passage of a group across the central meridian, and also at the time of the rapid development and decline of the corresponding solar phenomena. Thus, passage across the central meridian constitutes only one of the possible positions in which the effect of the active solar formation can occur (it would be more correct to say, of course, the radiation associated with it) in the earth's atmosphere.

There are certain other positions and states of an active center besides its position on the central meridian when it is capable of exerting an effect on the troposphere. It is interesting to note that the relationship of magnetic storms to the position of sunspot groups at the limbs of the disk was noted as early as 1907 by Cirera and Balcells (Ref. 29). The observations of Cirera and Balcells are interesting in that magnetic storms unquestionably are related to solar activity. Memery assumed that solar-induced processes can occur in the troposphere when an active region is situated at the limbs of the solar disk. However, if Cirera and Balcells had not drawn their conclusion, there would be considerable basis for assuming Memery's result to be random.

Also associated with the work of the Tashkent geophysicists is their working hypothesis of the influence of solar activity on meteorological phenomena. This so-called condensation hypothesis essentially involves the following: it is assumed that intensified short-wave and also corpuscular radiation, intensified during the irradiation of the earth by an active center on the sun, lead to an intensified dissociation of ozone, but at the same time dissociation of oxygen, thus saturating the atmosphere at the level of the maximum of the ozone layer with the free oxygen atom. The combination of the latter with the nitrogen molecule leads to the formation in the stratosphere of various oxides of nitrogen, especially nitric anhydride, the molecule of which can serve as a condensation nucleus for water vapor.

Cyclones, whose effect, as is well known, is propagated also into the stratosphere, draw stratospheric air from these layers into the troposphere, and the latter thus is enriched with condensation nuclei. This leads to the resolution of wet-unstable states which are typical for the low-latitude troposphere. The energy set free at the time of resolution of wet-unstable states also is a source of those intensifications of atmospheric circulation which are typical of the effect of solar activity on the troposphere (at least in the low geographic latitudes).

At this point we will not discuss the condensation hypothesis in detail, since this has been treated in the repeatedly cited book (Ref. 1). The condensation hypothesis explained the rapid reactions of low-latitude centers of action of the atmosphere to solar activity, but without the introduction of additional concepts it was difficult to understand how solar activity can influence the high latitudes. In order to explain this phenomenon the Tashkent geophysicists introduced a "series of circulation rings", but if this is correct, the lag in high-latitude meteorological phenomena relative to the most favorable date for the effect of an active solar center on the earth should be quite significant. This is the first, but by no means the only, and not the most important objection to the condensation hypothesis. In its first variant it was assumed that an intensification of solar activity leads to an

increase of the ozone content. Confirmation was discerned in the parallel behavior of the summary curves based on a 27.3-day calendar (for 1927) of ozone content and solar activity. However, in most cases the relationship between solar activity and ozone content is instead an inverse one.

In 1949, I. A. Prokof'yeva statistically demonstrated that solar activity first influences meteorological processes and the latter determine the ozone content over a particular area (Refs. 30, 31). The relationship between the ozone content in the atmosphere and synoptic conditions already has been known for a long time (Ref. 32). Thus, here is further demonstration of the untenability of the condensation hypothesis. Quite recently Willet (Ref. 33) succeeded in establishing a closer relationship between the total content of atmospheric ozone and solar activity than was assumed earlier. On the basis of data for two and one-half 11-year cycles it was possible to establish a quite close relationship between ozone content and the phase of the 11-year cycle.

A rigorous statistical evaluation of Willet's results, however, made by Mitchell (Ref. 34), revealed that they are of low reliability. It is entirely possible that this low reliability is associated with the method used in measurement of the total ozone content using a Dobson instrument, since the result of the measurements is dependent on atmospheric transparancy, which is extremely difficult to take into account.

Finally, it is impossible to pass by certain objections advanced by some meteorologists to the effect that condensation nuclei in the atmosphere usually are supplied abundantly from purely terrestrial sources, such as sea salt particles. It should be noted that an attempt has been made recently to revive the condensation hypothesis (Ref. 35).

Since we have touched upon the condensation hypothesis, which for some time was considered the most acceptable mechanism of the solar-tropospheric relationship, we will also mention another hypothesis which made its appearance during the period of transition from the third to the recent stages of study of solar-tropospheric relationships. We refer to the Haurwitz hypothesis (Ref. 36) concerning the heating of the ozone layer at the time of intensified solar activity.

The essence of the mechanism postulated by Haurwitz is that with an intensification of solar activity, primarily at the time of flares, there is an intensification of radiation in the region of wavelengths strongly absorbed by ozone. This leads to a heating of the ozone layer, especially its uppermost part, and the development of a temperature and therefore pressure gradient. In this way an influence is exerted on the general circulation of the atmosphere. The Haurwitz model, however, has at least two weak points. First, there is no basis for assuming that at the time of flares or at the time when large bright flocculas are

present on the solar disk, there is any intensification of radiation in the Hartley band (2,300-3,200 Å). Second, even if there was an intensification of radiation at these wavelengths, the result would be a decrease in the ozone content as a result of its direct dissociation and as a result of the reaction $2O_3 \rightarrow 3O_2$, which is very dependent on temperature. This was noted for the first time by E. Vassy (Ref. 37).

It can also be added that Haurwitz also does not discuss the mechanism of stratospheric control, in other words, how the circulation effect at the level of the upper stratosphere, that is, at a height of about 40 km, can influence tropospheric movements. All these objections lead to a hypothesis not proposed by Haurwitz. These same objections also are valid against the modification of the Haurwitz hypothesis which was porposed by Vassy, that it is oxygen, not ozone that is heated, that the heating occurs in a higher layer of the atmosphere than in the Haurwitz model, and that radiation is intensified in the shorter wave part of the spectrum (λ 1,300-2,100 Å), not in the region of the Hartley band.

In the late 1930's a considerable number of studies on the investigation of the relationship between solar activity and tropospheric phenomena was also carried out in the USSR, at Pulkovo Observatory, by the author and M. S. Eygenson. The results of these studies have been published in numerous articles, and also in the monograph "Solar Activity and its Terrestrial Manifestations", and other sources (Ref. 38). At this point, therefore, we will not enter into detail concerning these investigations, and we will discuss them only in general outlines.

The point of departure is the problem of establishing the relationship not between any meteorological processes and individual manifestations of solar activity, expressed by some particular index, but between processes on the sun and in the earth's atmosphere. This is accomplished by the author's method of representations, constituting a further development of the method of superposing of epochs. With respect to tropospheric investigations, we consider the direct precursors of our work to be not only the investigations of the Tashkent geophysicists and the studies of Memery, but also the interesting results obtained by Aufsess (Ref. 20), showing that the Azores anticyclone has a tendency to produce northward-directed spurs and the separation of centers directed toward Europe in those cases in which active regions pass across the solar disk, and in which the spot-formation process has already ceased, but in which extensive facula fields are still observed.

At the same time, M. N. Gnevyshev has advanced the concept of "impulses" of solar activity (Ref. 39) in which, in particular, he laid the basis of a typical model of development of an active solar center, later developed in the studies of D'Azambuja and Kiepenheuer (Ref. 40). We have striven to establish what geophysical, specifically, what

macrosynoptic processes are characteristic of a particular stage in the development of an active center, or to use the terminology accepted at that time, the phase of the "impulse" of solar activity. In this sense the comparison of solar and meteorological phenomena was raised to a new level. Whereas the investigations of the Tashkent geophysicists Memery and Aufsess can be characterized as a comparison of macrosynoptic processes and elements of solar activity, we endeavored to find in the development of the solar process the typical place occupied by some particular macrosynoptic phenomenon.

In the third stage of development of the sun-troposphere problem there was the first appearance of comparisons of variations of climate with solar cycles of greater duration than the 11-year cycle which merit attention. We might mention the work of Hanzlick showing that in a number of cases the 22-year cycle is expressed in the troposphere even more clearly than the 11-year cycle (Ref. 41). Of course. special attention should be given to studies which for the first time pointed out the role of the 80-90-year cycle of solar activity in the entire sun-troposphere problem.

In 1928 Easton investigated a large series of winters in Western Europe. An 89-year frequency of recurrence of severe winters was established (Ref. 42). This cycle also was confirmed by Koppen (Ref. 33). In 1936 Scherhag (Ref. 44) discussed the almost 100-year duration of the cycle of recurrence of severe winters in Europe; he agreed later (Ref. 45) that this cycle averages 89 years. It can therefore be stated that the 80-90-year cycle of solar activity, which, as we have already mentioned, was pointed out earlier by R. Wolf and later forgotten, was discovered for the first time in climatic phenomena at the transition between the 1920's and 1930's, and its presence on the sun was then confirmed by Gleissberg (Ref. 46).

The discovery of the long-term cycle in tropospheric processes, the famed "warming of the Arctic", attaining its maximum in the 1930's, and other phenomena, forced specialists to postulate that solar activity has an effect on climatic variations, and theoreticians were faced with the problem of determining the mechanism by which long-term variations of solar activity influence climate. In this respect an article published by Angstrom in 1939 (Ref. 47) is of great interest. Angström develops and supplements the important considerations concerning the nature of climatic variations expressed by Defant as early as 1921 (Ref. 48).

Defant was the first to point out the role of horizontal macroturbulent exchange and its variations as an important climatic factor and as the principal element responsible for climatic variations. It is known that heat transport in the troposphere occurs either by means of advection or by macroturbulent exchange. Advection transports heat along flow lines while macroturbulent exchange transports heat

perpendicular to them. The relationship between advective and macro-turbulent propagation of heat was investigated in detail by L. R. Rakipova (Ref. 49). Defant and Angstrom considered only macroturbulent heat exchange. The quantity of heat transported in this way is described by the formula which already has been cited:

$$Q = -c_p A \frac{1}{r} \frac{\partial T}{\partial \varphi}. \qquad (5.1)$$

Representing A (in the Schmidt form)

$$A = A_{\uparrow} + A_{\downarrow} = \frac{1}{st} \left(\sum_{s,t} m_{\uparrow} l_{\uparrow} + \sum_{s,t} m_{\downarrow} l_{\downarrow} \right), \qquad (5.2)$$

we have

$$Q = -c_p \frac{1}{st} \left(\sum_{s,t} m_{\uparrow} l_{\uparrow} + \sum_{s,t} m_{\downarrow} l_{\downarrow} \right) \frac{1}{r} \cdot \frac{\partial T}{\partial \varphi}. \qquad (5.3)$$

In the computations associated with general circulation, the area used is the part of a circle of latitude lying between the ocean level and the tropopause. The values $m_{\uparrow}$ and $m_{\downarrow}$ here should be considered as the masses of eddies, that is, in the case of an anticyclone as the excess of the mass of air above the mass which would be present at mean pressure, and in the case of a cyclone—as the deficit of the mass of air in comparison with mean pressure. As the mixing length $l_{\uparrow}$ and $l_{\downarrow}$ we can use very approximately the transverse dimensions of the eddies. Formula (5.3) shows to what a strong degree Q is dependent on the mass of the eddies (that is, on the intensity of the anticyclones and the depth of cyclones). From the law of accentuation of pressure fields under the influence of solar activity it therefore follows that the increase of the latter leads to an intensification of exchange. According to Angstrom, it is this which explains the character of recent climatic variation: intensified exchange directed warm air masses into the Arctic and transported a great quantity of cold masses from there. Since the area of the Arctic is less than the area of the temperate latitudes the temperature increase in the Arctic was more pronounced than the temperature decrease in the subtropics where this effect was expressed weakly.

The above-described mechanism whereby solar activity influences climate is the most probable, although the model given by Angström is extremely simplified. In particular, it follows from the monograph of L. R. Rakipova that it is necessary to take into account not only the

value Q, obtained using formula (5.3), but it also is necessary to take transport by advection into account, and it is entirely probable that advection along a meridian and advection along a parallel should be taken into account separately. Rakipova's expression for advection includes the circulation index, and this is dependent on the pressure gradient between the high and low latitudes; in turn this is associated with pressure formations, that is, the depth of cyclones and the intensity of anticyclones.

The discovery of very clear manifestations of long-term variations of solar activity in climate makes it essential to investigate the problem of the relative role played in climatic phenomena by solar cycles of different duration. As was pointed out in Chapter 1, the amplitude of the 11-year cycle considerably exceeds the amplitude of the 80-90-year cycle according to solar indices. Nevertheless, the climatic effects associated with the latter are considerably greater than those observed in the first.

Whereas the variations of climate associated with the 11-year cycle can be judged only from variations of the temperature of the tropics, which are very insignificant, and effects on precipitation and possibly evaporation, the 80-90-year cycle, through the mechanism of atmospheric circulation, causes extremely significant climatic variations. As we will see from the next section, there is basis for assuming that still longer cycles of solar activity cause even more significant climatic variations, although the amplitudes of these solar cycles are less than in the 80-90-year cycle. No satisfactory theory explaining this situation was proposed either in the third stage of study of solar-tropospheric relationships or later.

Section 4. Development of Investigations on the Sun-Troposphere Problem in the Recent Stage

In proceeding with a description of the recent stage of investigations of solar-tropospheric relationships, it should be noted in particular that one of its most characteristic features is an extraordinarily strong development of that group of investigations which we have termed "individual-statistical". The beginning was marked by the work of B. and H. Duell, published in 1948 (Ref. 50).

The Duells formulated the problem of discovering a relationship between the variation of surface atmospheric pressure and the solar corpuscular effect in particular. The index used for this solar influence was geomagnetic data, especially the five international quiet days and five international disturbed days for each month during the period 1906-1937. These days were used as points of reference. The data were

broken down by years with high solar activity, when the Wolf number exceeded 40, and years with low activity, when this number was $\leqslant 40$.

The data were also broken down by seasons: winter (November, December, January, February); spring and autumn together (March, April, September, October); and, summer (May, June, July, August). The initial investigations, based on data for Potsdam and Stikkisholm stations, revealed that clear results are obtained (a) only for years with mean Wolf numbers $\leqslant 40$ and (b) only for the winter months, from November through February inclusive. Thereafter only years with low solar activity and the winter months were used. The method of superposing of epochs was employed. By applying this method to data on atmospheric pressure for Potsdam and Stikkisholm (Iceland) for years with low solar activity, specifically 1910-1914, 1920-1924 and 1930-1935 (a total of 16 years) and to the winter season (from November through February inclusive), that is by using 320 reference days of each kind, the Duells obtained the data shown in Figure 42.

In Figure 42 data are shown not only for the two stations mentioned, but also De Bilt (Netherlands), Karlsruhe, Vienna, Wroclaw, Kaliningrad, L'vov and Kiev. The days considered were from -3 (that is, three days before a geomagnetic disturbance) to +11 (that is, 11 days after the reference day). For all the stations considered there is an opposite variation of pressure near geomagnetically disturbed and geomagnetically quiet days (solid and dashed curves, respectively).

Pressure behaves in the following way: after a particularly geomagnetically disturbed day the pressure drops and attains a minimum three days after the reference date. However, when a reference day is followed by a particularly geomagnetically quiet day, the pressure increases, also attaining a maximum three or four days later. It is on such days that the difference between pressure reckoned from a magnetically quiet and magnetically disturbed day attains a maximum value. The difference attains 2.6-3.2 mb (millibar).

In order to be able to represent the result in the form of a synoptic chart, the number of investigated stations was increased to 26. In addition to the stations mentioned above, data were included for Varda (Norway), Haparanda (Sweden), Trondheim, Lerwick, Oslo, Leningrad, Stockholm, Moscow and Copenhagen. Charts were compiled showing the deviations of pressure from the long-term norm for the corresponding months and for days before and after a reference day—from -1 to +8.

On day -1 from a particularly disturbed day, there were no appreciable deviations in the distribution of pressure from the mean pattern. The maximum difference in pressure within the chart is only 1.2 mb. Neither are there great differences of pressure on the disturbed day itself when the maximum difference is 1.6 mb. A sharp difference sets in

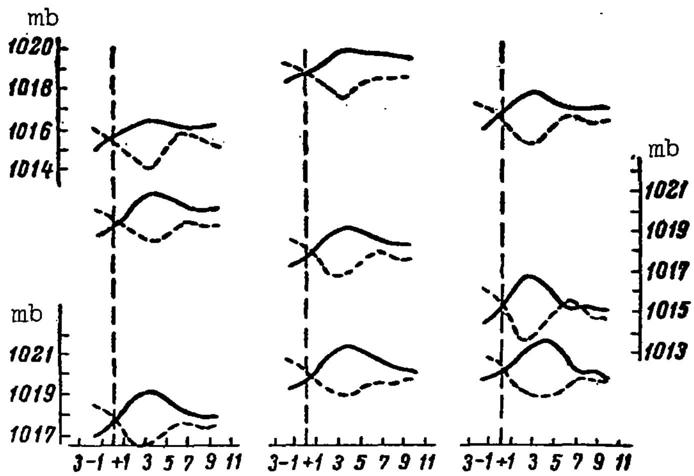


Figure 42. Mean variation of atmospheric pressure near geomagnetically disturbed (solid curves) and geomagnetically quiet (dashed curves) days for various meteorological stations in Europe (according to B. and H. Duell)

on the day following the disturbed day: the maximum difference in comparison with the mean chart already is 3.7 mb. Here the Duells apparently succeeded in noting the manifestation of the law of accentuation of pressure fields in an elementary macrosynoptic process.

Whereas on the mean chart the characteristic direction of pressure gradients is from southeast and south to northwest, on day +1 (after the geomagnetically disturbed day), the direction is from northwest to southeast. This confirms that a magnetic disturbance facilitates an Arctic intrusion into the European synoptic region. A region of high pressure is situated over Iceland and a region of low pressure lies over the Gulf of Bothnia. On the following day the gradient oriented from northwest to southeast is intensified still further.

The region of anomalies is displaced in the direction of the gradient, that is, toward the southeast. In general, this day is characterized by the maximum anomaly relative to the mean pattern (the maximum discrepancy between the "neighborhood" of geomagnetically disturbed and geomagnetically quiet days, as already mentioned, is observed somewhat later, specifically, three or four days after the reference day). Thereafter, the corpuscular effect begins to weaken, but only on the

eight day after the reference day does the situation return to approximately that which prevailed before the corpuscular effect began. The mean macrosynoptic picture at the time of corpuscular effects in an epoch of low solar activity and in the winter season, as determined by the Duells, corresponds to that which was discovered by later investigators, especially Koppe (Ref. 51). It therefore should be considered typical.

In the same study the Duells also investigated the effect of "ultraviolet influences" associated with flares. In these cases the reference days used were those when flares of importance 2-3 or 3 were observed between 0900 and 1500 ut. Only those cases were selected which within a 5-day period were not preceded by flares of equivalent importance. The data used by Duells in these cases covered the period from January 1, 1936 through December 31, 1941. The Duells selected 51 days with flares of the mentioned intensity, and by using the method of superposing of epochs obtained the distribution of atmospheric pressure within the range from one day before the reference day to eleven days afterwards. The following stations where used: Hamburg, Frankfurt-am-Main and Vienna.

A clearly expressed pressure maximum was observed at each of these stations four to six days after the "ultraviolet intrusion". In these cases the winter and summer curves are similar, that is, there is no appreciable difference between the behavior of pressure in April-September, on the one hand, and in October-March on the other. However, the amplitude of the pressure curves for the winter half-year is somewhat greater; its amplitude is 3.5 mb versus 2.3 mb in the summer months. The day-to-day variability of pressure attains a maximum on the second to fourth day after a strong ultraviolet intrusion. The Duells not only considered surface pressure, but also the absolute pressure pattern for the 500-mb level, which gave still more definite results.

On the day of occurrence of the "ultraviolet intrusion", the pattern in both summer and in winter is rather ambiguous, but on the next day it already is quite clear: a region of low pressure develops over Western Europe, and the anticyclone over Northern Europe is intensified. We should also mention the use by the Duells of an "inverse reference method", somewhat resembling the simplified variation of the "reflection method" introduced by the author in 1939. Here the reference days selected were those when the day-to-day variability of pressure at Frankfurt-am-Main station was not less than 5 mb. A total of 101 days with a pressure decrease, and 121 days with a pressure increase were analyzed for the period 1936-1941.

In addition to the reference day, the analysis included the days -11 to +6 inclusive. The characteristic numbers of calcium flocculas (an index which now has gone out of use) were processed. The curves of

the values of this index reveal an opposite trend in relation to the pressure curves, related to the dates of flares, as might be expected when using such a method; this shows the nonrandom character of the results when the dates of flares are used as the base. Pressure changes occur 3-5 days after change of the calcium flocculus index; an increase of this index is accompanied by an increase of atmospheric pressure and its decrease by a drop in pressure.

We have discussed the work of the Duells in detail because it seemingly served as a stimulus for an entire series of similar investigations in this area. Willett (Ref. 52) recently described investigation of the Duells as "clasic". However, it should be noted that their study really does not contain too much that is new. They developed the idea of a helioreference date, introduced into the sun-troposphere problem by prewar studies of Soviet investigators, and well-known processing methods are used. The service performed by the Duells was the synoptic analysis of such an important index as atmospheric pressure, compared with solar activity.

One of the weaknesses of this study is the absence of satisfactory statistical evaluations of the reality of the results. This appreciable gap was filled in an investigation made by Craig (Ref. 53). He also considered magnetically disturbed and magnetically quiet days (selected international days) for 1906-1939, but as a result of a more rigorous selection of years with low solar activity, he finally was left with only eight years.

The study was made for the four winter months in each of the selected years. Since five geomagnetically disturbed and five geomagnetically quiet days were selected in each month, Craig analyzed only 160 reference dates of each type. In addition to the reference dates he also considered days up to +10. Data on pressure were available for a large number of stations in the northern hemisphere between latitudes of 30° and 70°. Craig, like the Duells, obtains a negative correlation between atmospheric pressure near geomagnetically disturbed and geomagnetically quiet days.

The following statistical method was used for evaluation of the results. If at a particular station and for a particular value of the lag the pressure value for a disturbed day is greater than the mean for a particular season and level of solar activity, and the corresponding value for a quiet day is less, or vice versa, then some value X, used in the Craig method, is assigned the value 1. If both pressure values are above or below the mean, X acquires the value 0. After making such evaluations X is averaged for all geographic points. Craig obtained $X = 0.57$ for the reference day, but for the second day it already is $X = 0.71$, and attains a maximum on the fifth day when $X = 0.75$. Thereafter, X decreases. These means were obtained from 2,000 values.

Craig raises the following problems: (1) whether the used selection method introduces a correlation between pairs, and (2) whether the same values X can be obtained in a random sampling of the same data, that is, in a case in which a relationship between magnetic activity and pressure at sea level is known to be absent. Craig evaluates the value of the fictitious correlation that is introduced by the method itself, and finds that it does not exceed 0.6. The second problem is solved by comparing specific values for the second day after the reference day, i.e., 0.71, with the theoretical value, 0.61.

In this comparison it is taken into account how many standard deviations fall between these values. When computing the standard deviations it is necessary to take into account that not all of the pairs actually will be independent, because (a) pressure on the following days is determined to a certain degree by the pressure on the preceding days, and (b) because pressure at a certain station is not completely independent of pressure at other stations. Taking both of these circumstances into account, Craig obtains a final standard deviation of 0.017. Thus, six standard deviations fall between 0.71 and 0.61, that is, it is extremely improbable that one will obtain such a result in a random sampling of the same size.

The relationship between atmospheric pressure and geomagnetic activity is therefore confirmed statistically. Craig concludes that after a geomagnetic disturbance pressure increases in the high latitudes and decreases in the low latitudes. He also investigated the relationship between pressure and solar flares. His conclusions on this point are similar to those drawn by the Duells.

Also included among the investigations of this type is a study by Flohn (Ref. 54). He compared pressure anomalies for Oslo for the three winter months for the period 1884-1948 with magnetic activity. The latter was considered high when the international geomagnetic index was $C_i \geqslant 5$ (on a 10-unit scale). When there were from one to four such days in a month, the anomaly was negative, and its mean value was -1.4 mb. When the number of such days with this characteristic was from 5 to 11, the sign of the anomaly changed, and was +0.2 mb. When the number of days with the characteristic $C_i \geqslant 5$ was from 12 to 16, the anomaly was +1.6 mb.

Thus, with an increase of magnetic activity in the winter months at Oslo there is also an increase of the positive pressure anomaly. However, these results are not very reliable statistically. Flohn was in a position to draw more definite conclusions when he used thunderstorm activity instead of atmospheric pressure. In this case the study was for the summer months. The initial study was for the years 1923-1925 and 1930-1935.

Thunderstorm activity, analyzed for individual European stations and also for the regions of the Alps and Central and Northern Europe, was compared with the passage of "active spot groups" across the sun's central meridian, that is, spot groups in which there were flares or bright flocculas, and in which the growth or decay of the formations was quite turbulent. In most cases these were large spot groups, quite elongated longitudinally, which explains the necessity for limiting the investigation solely to years with relatively low solar activity in that at other times there is no possibility of separating the influence of one active region from the influence of others. In this case the statistical reliability of the result was higher than when comparing atmospheric pressure and geomagnetic activity.

In general, Flohn found that thunderstorm activity intensifies on the reference day itself, that is, on the day when an active region passes across the sun's central meridian. Flohn therefore concludes that thunderstorm activity is influenced by active solar wave radiation. Then, analyzing the five most geomagnetically disturbed days during the summer months of 1934-1938, and using them as points of reference (there were 125 such days during the period considered), Flohn discovered that thunderstorms are concentrated on the day directly preceding the reference day.

A similar investigation was made by Flohn on the basis of far more extensive data for 1884-1939. He considered days when geomagnetic activity on the 10-unit scale was 5 or more, but with the provision that within ± 4 days from the day of concern there were no days with a geomagnetic index $\geqslant 4$. He analyzed separately the following phases of the 11-year solar cycle: years near the minimum, years on the ascending branch, and years on the descending branch. The epoch of the maxima of the 11-year cycles was excluded from consideration due to an excessively high frequency of days with high geomagnetic activity.

In all Flohn considered 18 years near epochs of minima, 16 years on the ascending branch, and 22 years on the descending branch of 11-year cycles. The data on thunderstorms were for German and Dutch stations. The results indicate that thunderstorms occur predominantly three days before the reference day, that is, prior to a geomagnetically disturbed day. A verification of the reality of this conclusion by use of the X^2 test confirmed the nonrandom character of this relationship.

Flohn, however, took into account the absence of a nondependence of the observations, which is to be expected when the meteorological network is as dense as it is in Germany. Flohn attempted to overcome this difficulty by use of an inverse reference. He used as reference days those characterized by an intensification of thunderstorm activity, and the indices of geomagnetic activity were used for averaging. In this

analysis he used as reference days only those when thunderstorm activity was observed at not less than 50 percent of all stations in Northern Germany (the total number of stations in this zone was 1,300). It was found that magnetically disturbed days are concentrated on the second and third days after intensification of thunderstorm activity. This result confirms the conclusions made by Flohn on the basis of preliminary data.

The somewhat greater lag of geomagnetic disturbances relative to solar activity, obtained here in comparison with the initial results, which applied to the ascending branch of the 11-year cycle No. 17, apparently can be attributed to the fact that the later study also included other phases of 11-year cycles, and not only No. 17. Flohn's conclusion that thunderstorm activity is associated with solar active wave radiation, therefore, can be considered confirmed. However, the novelty of the conclusion drawn by this author should not be exaggerated. The writings of Septer and Brooks (Refs. 55, 14) already had indicated that the correlation coefficient between mean annual Wolf numbers and the number of days with thunderstorms is very high, and it is well known that Wolf numbers are basically an index of solar activity in relation to its variable wave radiation.

We will now discuss the work of L. A. Vitel's (Ref. 56) who used data for 1900-1939 from the synoptic catalog for Europe and for 1926-1939 for the Arctic. He selected 56 very large magnetic storms for 1900-1939, and 20 such storms for 1926-1939. He concluded that cyclonic circulation in the neighborhood of Iceland reveals two maxima within the considered period before and after a storm (from -3 to +3 days): the smaller is 1 or 2 days before the storm, and the other a day after it, being the principal maximum.

For the most part, cyclonic and anticyclonic circulation behave oppositely in all regions. In the Arctic, anticyclonic circulation attains a maximum one or two days before a storm, and cyclonic activity has a maximum three days after its commencement. In the Azores region the development of cyclonic and anticyclonic circulation is virtually the same. Over the southern part of the European USSR there are two peaks of anticyclonic circulation: one a day before a storm, and the other two or three days after it. Finally, the index of Atlantic circulation, constituting the sum of the mean depth of the Icelandic low and the mean intensity of the Azores anticyclone, increases sharply on the day after a storm (Figure 43).

We note that in the Vitel's catalog the entire natural synoptic region of Europe is broken down into eight regions, and for each of these on each day the synoptic field is characterized by one of the indices of an 11-unit scale from the index 9, corresponding to a well-developed cyclone, to the index 0, representing a well-developed

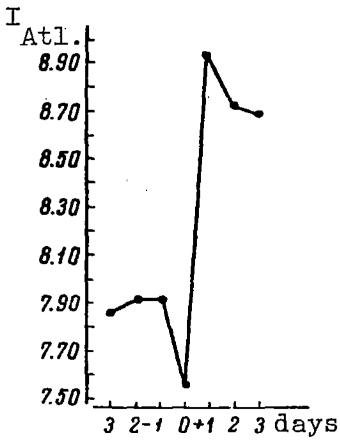


Figure 43. Behavior of the index of Atlantic circulation near the day with a magnetic storm (according to L. A. Vitel's)

anticyclone (11 is the neutral point index). It is superfluous to add that in this investigation, made by the method of superposing of epochs, the day of a magnetic storm is used as the reference day; the study was made in the range from the -3 to +3 days.

In addition to these investigations, which are the most typical of this aspect of the sun-troposphere problem in the recent period, a great many investigations of a similar character have been made and are being made (Refs. 57-69). The results of these studies sometimes are somewhat contradictory. In a number of cases their authors fail to find a relationship between solar and meteorological phenomena in the particular processes which they have subjected to investigation, and taking a skeptical viewpoint they draw the conclusion that there are no solar-meteorological relationships at all.

It also happens that a particular meteorological phenomenon undoubtedly is associated with solar activity, but the investigator has failed to select the proper geophysical index characterizing it, or the proper data with respect to season or phase of the 11-year cycle, without taking into account long-term variations of solar activity, or, finally, has used an unsuitable solar index and obtains ambiguous results, and in certain cases even an apparently negative conclusion. However, in most cases relationships are established as a result of these investigations.

The author also made a number of investigations in this field of the sun-troposphere problem in 1957-1961. Initially they were similar to the work of the Duells, but differed from it in that the frequency of occurrence of circulatory mechanisms as defined by Dzerdzeyevskiy was used in place of atmospheric pressure (Ref. 70). The epoch analyzed was 1933-1937, that is, the ascending branch of the 11-year cycle No. 17. As in the work of the Duells the reference days used were geomagnetically disturbed and geomagnetically quiet days (selected international days, five days of each kind each month). For the five years, 300 disturbed and an equal number of quiet days were selected. Distribution curves of circulatory mechanisms were constructed for both types of days. The study was based on the 13 principal types of circulatory mechanisms without additional division into subtypes.

Since a certain number of days in the year are characterized by circulation that does not conform to any of the 13 types, it was necessary to introduce a fourteenth value denoted y, i.e., atypical days. Since the frequency of occurrence of the different types of circulatory mechanisms is different for purely dynamic-climatological reasons, a distribution curve of days with different types of circulatory mechanisms was constructed for the five years, bearing no relation to magnetic disturbance. In other words, the frequency of recurrence of Dzerdzeyevskiy circulatory mechanisms was clarified. This distribution then was broken down into the distributions for magnetically disturbed and magnetically quiet days.

Such normalized distribution curves apparently were free of the influence of nonuniform frequency of days with different circulatory mechanisms. These normalized distributions were determined both for the reference day itself and for 12 other days: six preceding and six following the reference day. The days preceding the reference day usually are denoted by a minus sign and the days following by a plus sign. Thirteen normalized distribution curves were obtained in this way.

Data for the zero (reference) day and for the +1 day are represented graphically in Figures 44 and 45. The circulation types are plotted along the x-axis, and the relative frequency of the particular type along the y-axis. The graph shows that the distribution curve whose reference day is a geomagnetically disturbed day (solid curve) and the distribution curve whose reference day is a geomagnetically quiet day (dashed curve) reveal a clearly expressed opposite behavior. The correlation coefficient between the curves is -0.83 for the zero day. However, it must be remembered that when 5 days are selected initially from the 30 days of a month, a sample consisting of five arbitrarily chosen days from the 25 days remaining after the first five-day sample, selected on the basis of a definite criterion, should be related to the latter by some a priori negative correlation coefficient.

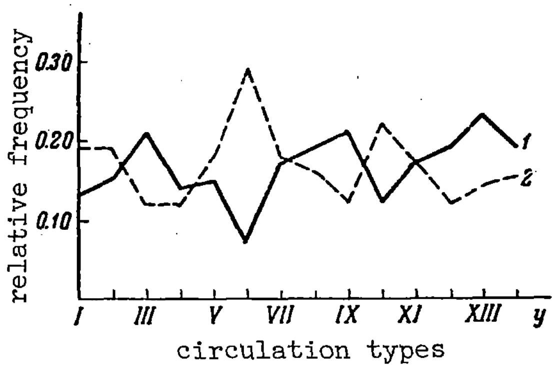


Figure 44. Relative frequency of circulatory mechanisms as defined by B. L. Dzerdzeyevskiy for geomagnetically disturbed (1) and geomagnetically quiet (2) days

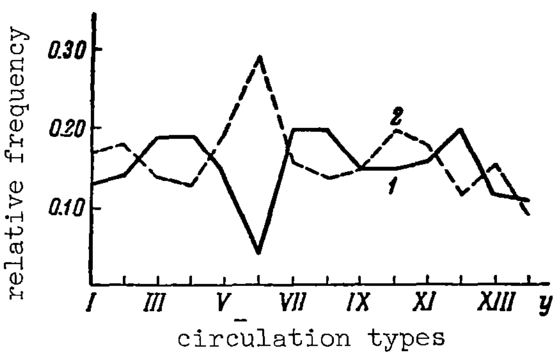


Figure 45. Relative frequency of circulatory mechanisms as defined by B. L. Dzerdzeyevskiy a day after a geomagnetically disturbed (1) and a geomagnetically quiet (2) day

The very fact of the separation out of five days characterized by some definite criterion from the initial population in each month lowers the probability of finding any criterion in the remaining 25-day population which will characterize the five days initially selected. Thus, in order to evaluate correctly the correlation coefficients obtained in this analysis, it is necessary to compare them with an a priori correlation coefficient. The comparison can be made using Fisher's z (Ref. 71).

The value of the a priori correlation coefficient is dependent on the ratio of the size of the initial sample to the total size of the population. It can be shown that when there are five initially selected

days from a total population of 30 days in each month, the a priori negative correlation coefficient is -0.33 (Ref. 72). A comparison of the above-derived correlation coefficient -0.83 with the a priori coefficient shows the reality of the first.

In the further investigation of the problem it was of interest to clarify whether in actuality the maximum discrepancy of the curves of relative frequencies of circulation types near geomagnetically quiet and geomagnetically disturbed days is observed for an index closer to the reference date than that used by the Duells, that is, when the tropospheric index is atmospheric pressure. An investigation, therefore, was made similar to that in Ref. 70, but using data for different years and seasons than used in the investigation of the Duells. The following years were used: 1909-1914, 1919-1924, and 1929-1935, and the months from November through February inclusive. A total of 320 magnetically disturbed days, and an equal number of magnetically quiet days were used. The period before and after these events considered was the -3 to +6 days.

The most clearly expressed opposite behavior of the curves whose reference days were geophysically disturbed and geomagnetically quiet days, respectively, is observed on the +2 day. On this day the correlation coefficient between the curves is -0.79. For the +3 day it is -0.72. A comparison with the a priori correlation coefficient by use of the z test shows that the value -0.79 is real and -0.72 is not.

The most clearly expressed opposite behavior in the distribution of the relative frequencies of the Dzerdzeyevskiy circulatory mechanisms near magnetically quiet and magnetically disturbed days is observed on the second day after the corresponding reference day. For this same epoch and season atmospheric pressure over Western Europe and the Eastern Atlantic reveals the most clearly expressed opposite behavior near geomagnetically disturbed and geomagnetically quiet days on the third day after the corresponding reference day. It therefore would be possible to conclude that corpuscular solar influence is manifested first in the direction of air flow and then in the pressure field (Ref. 73).

However, it is more precise to express this somewhat differently. It must be taken into account that the decisive parameter in the classification of Dzerdzeyevskiy circulatory mechanisms is the direction of Arctic intrusions. In this respect the result means that the corpuscular effect is manifested first in the troposphere of the high latitudes, and then, through the processes developing there, exerts an influence on the pressure regime of the entire remaining part of the hemisphere.

An analysis of the relative frequencies of days with various Dzerdzeyevskiy circulatory mechanisms near geomagnetically quiet and geomagnetically disturbed days for the solar activity cycles Nos. 14-18

shows considerable positive and negative anomalies of the relative frequencies of definite mechanisms. The following types of circulation have this property: sixth (Arctic intrusions into the Pacific Ocean); thirteenth (penetration of cyclones into the Central Arctic); and, first (zonal circulation). In addition, the properties of certain circulatory mechanisms can be concentrated near geomagnetically disturbed or geomagnetically quiet days in one epoch and not observed in another.

The above is illustrated in Table 31, which gives only the most clearly expressed positive or negative extrema. It follows from this table that the most clearly expressed concentration in dependence on the presence and absence of geomagnetic activity is revealed by circulatory mechanisms of types 13 and 1. Type 13 characterizes the penetration of cyclones into the Central Arctic and type 1 characterizes zonal circulation in the temperate latitudes. The most nonuniformly disturbed types relative to days of the presence or absence of solar corpuscular effects, therefore, are the circulatory mechanisms characterizing zonal circulation.

At the same time it should be noted that on the ascending branch of cycle No. 14 and in cycle No. 15, the behavior of type 13 circulatory mechanism is different. In the first case the relative frequency of this mechanism increases after a geomagnetically disturbed day and in the second it is maximum either on the geomagnetically quiet day itself or a day later. It therefore can be postulated that the appearance of corpuscular activity in the troposphere is dependent on a phase of a solar cycle of greater duration than 11 years. This possibly is the 22-year cycle or possibly the 80-90-year cycle.

We will now attempt to generalize somewhat the results of the discussed type of investigations on the sun-troposphere problem made during approximately the last decade. It first became clear that the manifestation of active wave and solar corpuscular radiation in meteorological processes is appreciably different. An ordinary case of the passage of a well-expressed active region across the sun's central meridian gives a complex effect. In the simplest case it can be assumed that a considerable exposure of the earth to active solar wave radiation occurs when the corresponding active region of the sun is situated near its central meridian. The corpuscular effect appears somewhat later and this lag is associated with the time required for propagation of the corpuscular stream.

As is well known, the magnetic storms which can be associated with this type of active center are storms with a sudden commencement. They also can occur when the active region is not on the central meridian. The effect of the wave radiation from an active center also can occur when the center is not near the center of the visible disk. However, the effectiveness of the effect increases when the center has a central

Table 31

Epoch	Day after event	Number of circulatory mechanism	Character of day	Type of extrema
1901-1905	+3	13	Disturbed	Maximum
1913-1917	0	13	Quiet	Maximum
	+1	13	Quiet	Maximum
1918-1922	+1	13	Quiet	Maximum
	+4	1	Quiet	Maximum

position. The effect of corpuscular radiation causing magnetic storms with a sudden commencement therefore precedes an intensification of active wave radiation, especially in the X-radiation region, causing a complex of phenomena known as a sudden ionospheric disturbance. Corpuscular effects in pure form, that is, those which are not preceded by appreciable intensifications of wave radiation, and which are the cause of development of type-M magnetic storms (that is, those associated with M-regions) very probably exert somewhat different influences on the troposphere than those which are associated with flares.

With respect to one of the regions most sensitive to variations of solar activity, the European synoptic region, the typical manifestation of complex solar effects develops in the following manner. On the day of a strong flare or a day later or at the time of passage of an active region with numerous bright flocculas across the sun's central meridian, the pressure in Central Europe drops and the temperature increases. The pressure gradient between the Azores and Iceland increases due to a pressure increase in the Azores anticyclonic system, and at the same time in the Icelandic low system (manifestation of the law of accentuation of the pressure field in an elementary process).

In the summer, at times of passage of an active region in which strong flares are observed across the sun's central meridian, and sometimes earlier (by approximately a day) there is an intensification of thunderstorm activity and some increase in the quantity of precipitation. Following the wave action the corpuscular effect first facilitates the development of cyclonic circulation in the region of Iceland and an increase of the index of Atlantic circulation, usually on the day following a strong magnetic storm. The increase of anticyclonic circulation in the Arctic 1 or 2 days before a magnetic storm forces us to assume that wave radiation in the high latitudes can facilitate anticyclogenesis.

Two days after a magnetic storm a cyclone of Icelandic origin is displaced eastward, leading to a pressure increase in the neighborhood of Iceland, and, therefore, to a decrease of the pressure gradient between Iceland and the Azores. As a result the zonal circulation weakens and meridional circulation sets in. On the days which follow, an Arctic intrusion develops in Europe on the rear of a cyclone moving eastward from Iceland. This intrusion is in the form of a primary or secondary anticyclonic formation, only shortly before arriving from the neighborhood of Iceland.

These processes apparently do not adhere to such a clear pattern, but it can be stated nevertheless that there also is an intensification of zonal circulation on days with strong flares which is replaced by meridional processes after magnetic storms associated with these flares. According to aerological data, in years of high solar activity, zonal circulation in this synoptic region is observed after geomagnetically quiet days, and meridional circulation is observed after geomagnetically disturbed days.

We will now return to recent investigations of the manifestation of the 27-day cycle of solar rotation in the lower layers of the earth's atmosphere. After the work of the Tashkent geophysicists we can note several studies made by Abbot, published in 1947-1949 (Refs. 74-76). Abbot concludes that there is a 27.0074-day cycle in the Washington area. The effort to obtain an excessive accuracy in determination of the duration of the cycle causes one to doubt the entire result obtained by Abbot. Of some interest in one of these studies is a statistical method of evaluating the reality of the results; the dispersion of the mean curve relative to individual curves, determined from the rows of the calendar, is compared with the dispersion of this sample from random numbers.

Several studies devoted to the manifestation of the solar rotation cycle in the troposphere have been made by L. A. Vitel's (Refs. 77, 78). In the first the author considered the more general problem of the manifestation of active solar longitudes on macrosynoptic processes. In particular, it was demonstrated that the frequency of a positive temperature anomaly at Leningrad in 1951, as determined using a 5-day moving average, revealed two maxima within the 27.3-day cycle. The first fell on the 9th and the second on the 27th day. The position of the maxima is well maintained.

Proceeding on the basis of these results, Vitel's constructed a table giving the pattern of solar rhythms. He then analyzed observations for the following stations: Kazan', Kiev, Kirov, Orenburg, Volgograd, Minsk, Odessa and Khar'kov and longer series (from 1890) for Leningrad and Moscow. He used the deviations of the temperature of the corresponding day of the year from the long-term mean to eliminate the annual

variation. These anomalies were converted into rank numbers (Ref. 79); the largest negative temperature deviation from the norm for a particular solar rotation was assigned rank number 1, and the largest positive deviation was assigned rank number 27. The remaining numbers constituted a range of values.

In a case in which there were two identical values of an anomaly in a particular 27.3-day cycle, both days were assigned the same rank number, equal to the mean which should correspond to the particular anomaly and the next rank number. The preliminary results of analysis of a three-year series of temperature anomalies for Leningrad revealed the opposite behavior of recurrence of rank numbers corresponding to cases of strong cold waves (Nos. 1-5) and considerable warmings (Nos. 23-27) in the solar rotation cycle. Vitel's constructed curves showing the recurrence of rank numbers for 1951-1955 for all the above-mentioned stations, and on this basis a summary graph of the recurrence of these rank numbers for all 10 stations. The results, smoothed using 3-day moving averages, are shown in Figure 46. There is a clearly expressed concentration of strong coolings on the 19th and 25th days of the solar calendar.

It is of still greater interest to analyze the day-to-day solar calendar distribution of the dates of transformations of the principal atmospheric circulation forms as classified by G. Ya. Vangengeym. These results for the period 1891-1956 are shown in Figure 47. The solid curve represents the number of cases of transformations of easterly and meridional types of circulation into westerly in the course of the 27.3-day cycle, the dotted curve represents the transformations of

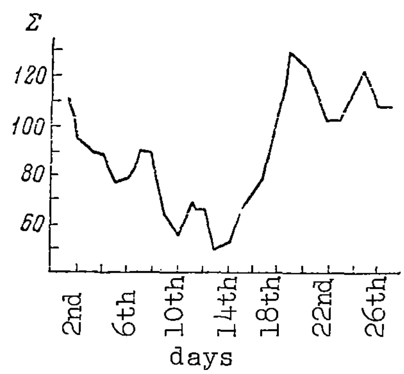


Figure 46. Frequency of rank numbers of negative temperature anomalies in the 27-day cycle during the period 1951-1955 (according to L. A. Vitel's)

easterly and westerly types into meridional, and the dashed curve the transformation of westerly and meridional types into easterly. It can be seen that transformations with a transition to westerly-zonal type of circulation occur most frequently near the 4th and 23rd days of the solar calendar.

Transformations with transition to a meridional type are observed most frequently near the 19th day. The latter also explains the cold waves causing a concentration of negative temperature anomalies near this day, as mentioned above. We have discussed the group of studies made by Vitel's in detail because they constitute the most basic result in this aspect of the sun-troposphere problem from among the investigations made in recent years.

Among the other studies which can be mentioned is one by Visser (Ref. 80), who made a harmonic analysis of the maximum daytime temperatures at a number of stations in the United States in the 27-day cycle in 1947 and 1953. Visser was able to show that the maximum temperatures are observed 13 days before the appearance of sunspots and minimum temperatures reveal a tendency to predominate the same number of days after appearance of spots. Rangardjean (Ref. 81) established a number of relationships between processes in the upper layers of the earth's atmosphere and the weather. A 27-day recurrence of meteorological phenomena was discovered clearly in 1913-1914, that is, in an epoch of a very deep minimum of solar activity. In 1952-1953 the atmospheric pressure in southern India changed in accordance with a 27-day cycle.

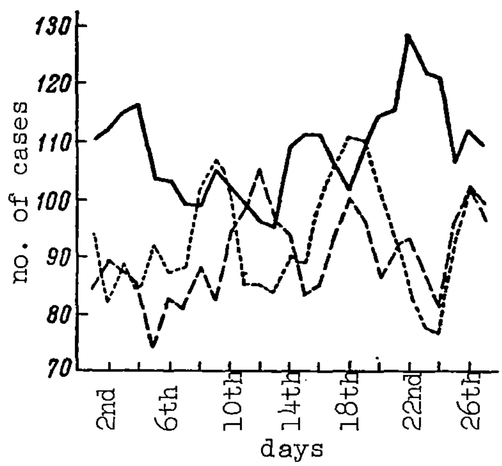


Figure 47. Frequency of transformations of circulation types in the 27-day cycle according to the G. Ya. Vangengeym classification (according to L. A. Vitel's)

In the case of high solar activity, such as in 1948-1949, a 27-day recurrence of atmospheric pressure values was not observed. Rangardjean attributes this to the presence of an excessively large number of active centers on the sun. Thus, a brief summary of the study of the manifestation of the 27.3-day cycle in the lower layers of the atmosphere would be as follows. This cycle is revealed both by a number of meteorological elements (pressure, temperature anomalies) and circulation characteristics. In certain years this "rhythm of position" is manifested more clearly than in others. It would appear that in epochs of high solar activity it is not always possible to discover the solar rotation cycle in meteorological phenomena. During recent decades it has been possible to note quite definite days on the solar calendar when synchronous weather changes occur over an extensive area. The factor responsible for this is the occurrence of rather sharp changes of the type of atmospheric circulation at such times.

We will now discuss recent investigations of manifestation of the 80-90-year solar cycle in the lower layers of the earth's atmosphere. Development along this line has been facilitated in recent years by two circumstances: (a) a number of direct investigations of this solar cycle were made, and (b) many investigators, for fully understandable reasons, have attempted to make a detailed study of the remarkable climatic anomaly that began with the transition from the 19th to the 20th century, and attained its apogee in the 1930's.

We will first discuss the work of L. A. Vitel's, published in 1948 (Ref. 82). Vitel's establishes a relationship between the intensity of atmospheric circulation in the European synoptic region and long-term variations of solar activity. He was able to show that there is an appreciable difference between circulation in the first two decades of the current century and the two that followed. The number of days with deep cyclones in three northern regions, the Greenland, Barents and Kara Seas, in the first 11-year cycle of our century (that is, in cycle No. 14) changes parallel with the variation of Wolf numbers, but beginning with cycle No. 15 it is possible to note a divergence of the corresponding curves, and with a transition from cycle No. 16 to cycle No. 17 the maximum number of cyclones begins to fall in epochs of sunspot minima.

In general, however, the number of deep cyclones and well-developed anticyclones increases with the phase of the 80-90-year cycle. Vitel's introduced an appreciable correction in the law of accentuation of pressure fields for the region considered. It was found that this law by no means always has such a simple form as formulated by its discoverers: Vize, Walker and Fedorov. An intensification of solar activity in the long-term cycle does not always lead simultaneously to a deepening of cyclones and an intensification of anticyclones. The development of solar activity can include rather prolonged phases characterized either by an intensification of anticyclones or a deepening of cyclones. The

first or second effects even can be transformed into the opposite effects, that is, there can be such periods when cyclones for the most part continue to deepen with an intensification of solar activity and at the same time anticyclones become weaker rather than intensifying.

Another situation also is possible; with an increase of solar activity, anticyclones intensify, but the cyclones fill. We can illustrate (Figures 48-50) the variation of 10-year moving averages of the percent of days with deep cyclones (solid curve) and well-developed anticyclones (dashed curve) according to Vitel's in the following three regions of a synoptic catalog which he compiled: Iceland (Figure 48), the Kara Sea (Figure 49), and the Eastern Atlantic to the north of the Azores (Figure 50). It follows in particular from this work by Vitel's that a relationship continues between the number of deep cyclones and solar activity even during rather high phases of development of the current 80-90-year solar cycle, provided the solar activity index used is $\bar{a}$, not Wolf numbers. It was this result which for the first time made it possible to evaluate the geophysical significance of the index $\bar{a}$, which can be seen most clearly by a comparison of $\bar{a}$ with the index of Atlantic circulation I_{Atl}, introduced by Vitel's and representing the annual sum of the number of days with deep cyclones in region No. 1, and well-developed anticyclones in region No. 5. The variation of the values I_{Atl}, smoothed using the three-term formula, and the variation of the transformed index $\bar{a}$ $A = (\bar{a} - 1)\ 1{,}000$ has been shown in Figure 51.

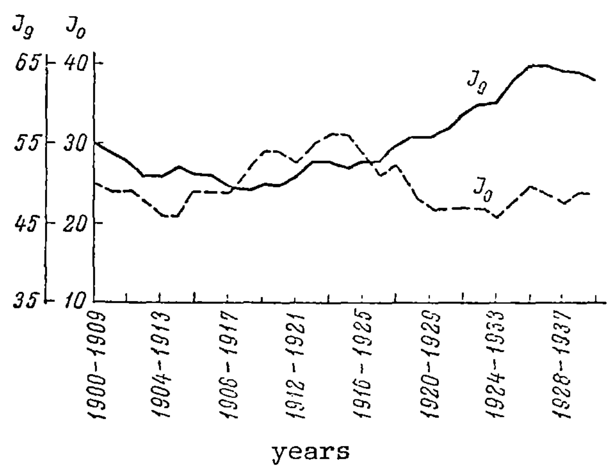


Figure 48. Well-developed cyclones and anticyclones in the neighborhood of Iceland (according to L. A. Vitel's)

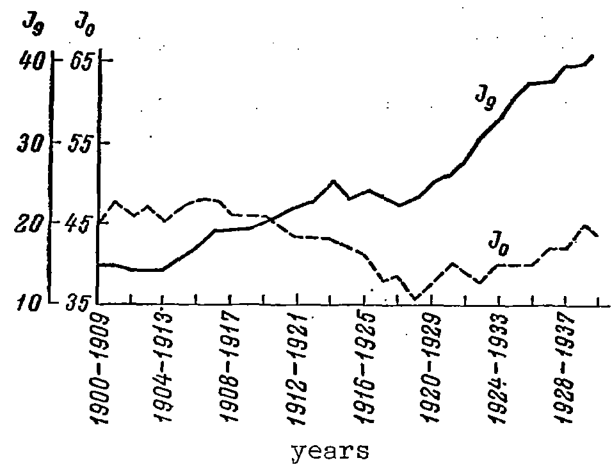


Figure 49. Deep cyclones and well-developed anticyclones in the neighborhood of the Kara Sea (according to L. A. Vitel's)

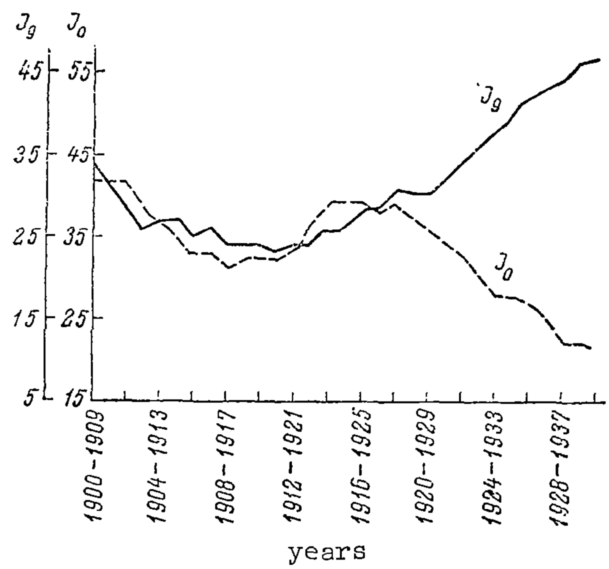


Figure 50. Deep cyclones and well-developed anticyclones in the neighborhood of the Azores (according to L. A. Vitel's)

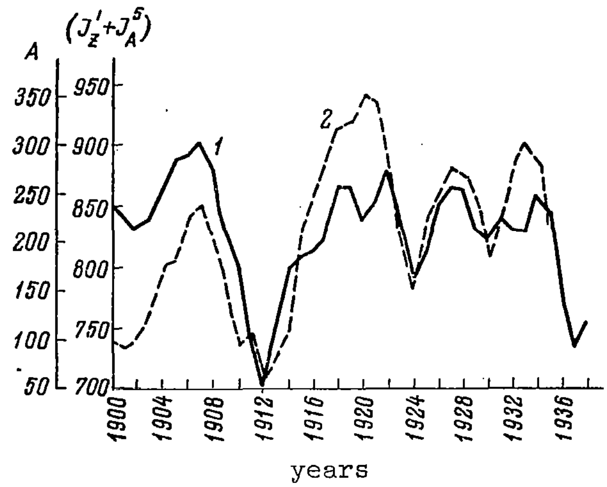


Figure 51. Moving 10-year mean indices of Atlantic circulation (1) and the transformed index $\bar{a}$ (2) (according to L. A. Vitel's)

The work of Vitel's clearly reveals the recent climatic variation on the basis of intensity of circulation in a limited area, but a region extremely representative for sun-earth relationships. This same climatic variation is also revealed clearly in the change of type of circulation. At this point we should mention the results obtained by A. Ya. Bezrukova (Refs. 83, 84). This author compared the 10-year moving sum of mean Wolf number values and the number of days annually on which a particular type of atmospheric circulation in the northern hemisphere (according to Dzerdzeyevskiy's classification) was observed. A particularly clear result was obtained with respect to the number of days with the third group of circulatory mechanisms (Arctic intrusions simultaneously from two or more directions). This result is illustrated in Figure 52. The solid curve gives the number of days with meridional circulation, and the dashed curves give the Wolf numbers (we repeat that in both cases they have been represented in the form of their 10-year moving sums).

Bezrukova also established the manifestation of the 80-90-year cycle in the levels of large lakes. With an increase of solar activity there is a decrease in the number of days with a meridional type of circulation and a simultaneous decrease in the level of one of the largest American lakes.

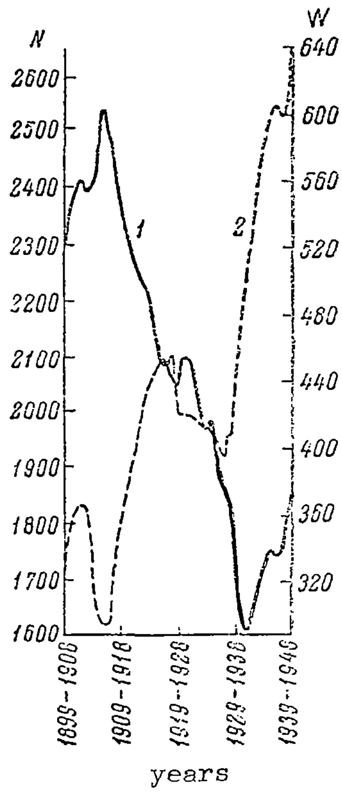


Figure 52. Moving 10-year sums of the numbers of days with meridional circulation (1) according to the B. L. Dzerdzeyevskiy classification and Wolf numbers (2) (according to A. Ya. Bezrukova)

L. A. Vitel's and A. A. Girs also investigated the manifestation of the 80-90-year solar cycle in the recurrence of changes of types of circulation as defined by G. Ya. Vangengeym (Refs. 85-87). Girs used the integral curve method and Vitel's the method of formation of moving averages or sums for 22- and 11-year periods. The principal conclusion drawn by Girs is that there are epochs each of which sometimes include two or three decades characterized by a predominance of one particular type of Vangengeym circulation. There is then a change of the circulation epoch. Figure 53, taken from a study by Vitel's (Ref. 85), gives the variation of the 10-year means of the number of transformations of Vangengeym circulation types.

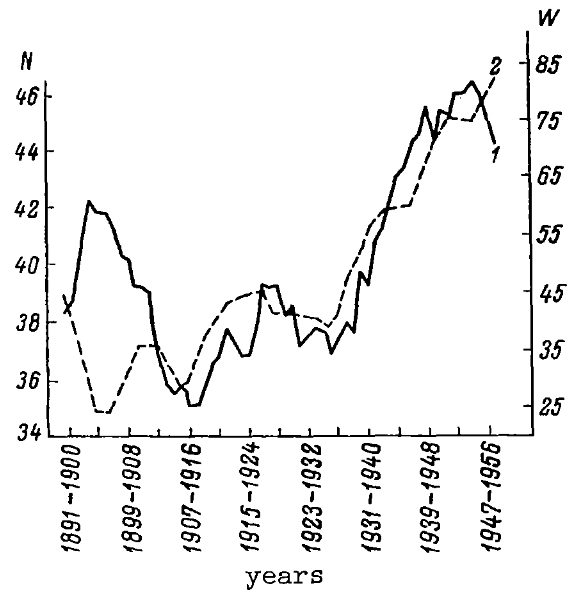


Figure 53. Moving 10-year mean numbers of transformed circulation types of G. Ya. Vangengeym (1) and Wolf numbers (2) (according to L. A. Vitel's)

The solid curve corresponds to the 10-year moving mean of the annual number of transformations of circulation forms and the dashed curve the moving mean of Wolf numbers. Among the other investigations of manifestations of the 80-90-year cycle of solar activity in the troposphere, we can mention the results obtained by N. V. Kolobov (Ref. 88). Using data for the period 1878-1958 he discovered a manifestation of this cycle in the precipitation at Kazan'. Before 1900 there was a continental period: a cold winter, warm summer and decrease in total annual precipitation. Beginning in 1900 the conditions changed.

Closely related to the studies of Vitel's and Girs are the results of a study of variations of atmospheric circulation over the North Atlantic obtained by B. L. Dzerdzeyevskiy (Ref. 89). The results obtained by Kolobov were confirmed in a study by A. I. Mishkarev (Ref. 90). He established that the integral curve of the sums of precipitation at Kazan' ties in well (direct dependence) with the frequency of Vangengeym's westerly type of circulation, and also with the meridional and easterly types, according to this same classification (inverse relationships).

Mishkarev noted that there was a decrease in the annual sums of precipitation beginning in 1936, that is, parallel with the long-term increase of Wolf numbers. An investigation made by D. A. Ped' and T. V. Sidochenko (Ref. 91) shows that the intensity of zonal circulation is dependent on solar activity. A comparison of the integral curves of the number of days with such a type of circulation and Wolf numbers for 1938-1957 makes it possible to discover a minimum of zonal circulation in 1938 and a maximum in 1956.

A number of investigations have been devoted to long-term variations of the hydrological regime. Horn (Ref. 92) points out that variations of sea level along the German coast have a long-term variation revealing a cycle close to 80 years. It is of considerable interest to study the series of investigations made by I. M. Soskin (Refs. 93-95). He discovered an approximately 80-year variation in ice conditions in the Barents Sea. Integral curves show that this difference is inverse, that is, the long-term increase in solar activity corresponds to a long-term decrease in severity of ice conditions and vice versa. It was also established that in periods of the long-term decline of solar activity, that is, near the epochs of the minima of the 80-year cycles, the level of the Caspian Sea is above the norm and vice versa.

The salinity and water exchange of the Baltic Sea also conform to long-term cyclic variations of solar origin. These effects are the result of changes in atmospheric circulation. N. V. Butorin (Ref. 96) has described a rise in the level of the Atlantic Ocean in connection with solar activity. Between the 1890's and the 1940's this rise was more than 6 cm. Until the mid-1930's the relationship to Wolf numbers was good.

Krapfenbauer (Ref. 97) compared the temperature conditions prevailing from 1901 to 1925. He also attributed the change in conditions to a "79-year" cycle. In this study there is an interesting attempt to trace the epoch of the minimum of such a cycle in the past, specifically in 1670, 1750 and 1820, and the epochs of the maximum in 1699, 1780 and 1850. Krapfenbauer's conclusions, therefore, are based not only on the last long-term cycle, but on others as well.

Therefore, if an attempt is made to summarize the results of studies made up to now of the manifestation of the 80-90-year cycle in atmospheric circulation, the following can be stated: in the first quarter of the long-term cycle the intensification of solar activity behaves in the spirit of the "classical" accentuation law. In the second quarter of the cycle the accentuation law is modified. The development of the 80-90-year cycle leads to a decrease in the number of days with Arctic intrusions from two or more directions simultaneously; at the same time, with development of the cycle there is an increase in the number of cases of transformation from one circulation type to another.

A sharp change of conditions sets in with a transition from the last phase of the old 80-90-year cycle to the first phase of the new one.

Whereas at the end of the cycle there are clearly expressed continental tendencies in climate, at the beginning of the next there is a rather rapid transition to marine tendencies. It can be said that in general the results of investigations of the variation of circulatory characteristics in the 80-90-year cycle confirm the theoretical model discussed above (the Defant-Angström model), but that in actuality the processes involved are considerably more complex than indicated by this model.

During the past 10 to 15 years attempts have been made to detect cycles of climatic variations longer than the 80-90-year cycle, but also of solar origin. One of the principal difficulties leading to the ambiguous character of the results is that, as pointed out in Chapter 1, we have virtually no data concerning solar cycles of such a duration. In attributing the detected climatic cycles to a solar origin, the basis used is that there are no intra-atmospheric factors that could lead to cycles of such a great duration.

It would be better, of course, to have cases for which there are additional data to support the hypothesis that the climatic cycles detected are in fact solar-induced. Mention, therefore, should be made of studies published by I. V. Maksimov (Refs. 98, 99). Using periodogram methods, he first discovered an 80-year cycle, and later a 600-year cycle, in sequoia growth rings. As pointed out in Chapter 1, there are data indicating that the 600-year cycle, in actuality, is of solar origin. We can also mention an article by Richter-Bernburg (Ref. 100), who discovered a 180-190-year cycle in the strata of certain deposits of Permian age. As stated in Chapter 1, cycles of such a duration apparently can be detected now on the basis of Wolf numbers (Ref. 101).

Link (Ref. 102) has indicated the existence of a 400-year cycle in climate which is related to a similar cycle of solar activity. His results are close to the conclusions of Rodewald (Ref. 103), who feels that the increase of temperature in the high and temperate-high latitudes observed in the first half of this century has restored the climatic conditions which were observed before a worsening of climate during the period 1540-1600.

It is Rodewald's opinion that the next few decades will be characterized by cooling and the advance of glaciers, and this is associated with that phase of the 80-90-year cycle into which we now are entering. Flohn made a study of the periods of climatic cycles beginning with A.D. 1000 and concluded that at least a majority of such cycles, if not all, have an "extraterrestrial" origin (Ref. 104). Fairbridge (Ref. 105) points out the presence of cycles of long duration in the level of

the world oceans. He found cycles with durations of 550, 1,100, 1,650 and 3,600 years. The cycle of 550 years is of particular interest because it is close in length to the 600-year cycle discovered on the sun (according to Schove's data in Ref. 106). This cycle in actuality is somewhat shorter than 600 years, and is closer to Fairbridge's hydrological cycle. We also should mention a number of cycles in the variations of the hydrological regime discovered by A. V. Shnitnikov (Ref. 107); however, it is impossible to be sure that the latter are all of solar nature.

With respect to cyclicity in paleoclimate, the most consistent views on this subject have been expressed by Willett and P. P. Predtechenskiy (Refs. 108, 109). Willett feels that both climatic variations and variations of solar activity are made up of a whole "set" (spectrum) of cycles of different scale. Willett distinguished three characteristic types of climatic conditions in the geological past: glacial, interglacial, and intermediate.

Glacial conditions are characterized by a moderate intensification of solar activity, for the most part with a predominance of ultraviolet radiation over corpuscular radiation. Interglacial conditions occur when there is an "absolute minimum" of solar activity. Finally, intermediate conditions are characterized by extraordinarily high solar activity with a predominance of active corpuscular radiation over active ultraviolet radiation.

Willett discerns a similarity in the development of extremely long solar cycles and cycles of far shorter duration, and even a similarity of long solar cycles and the development of individual active centers, in the sense of an alternating of phases of radiation of active shortwave and corpuscular solar radiation.

Predtechenskiy, like Willett, feels that it is possible to explain all paleoclimatic variations of extremely long duration on the basis of solar activity cycles. Predtechenskiy (Ref. 111) has formulated the problem of the patterns of alternation not only of major periods of glaciation (the period of 200-300 million years that is frequently mentioned in the literature, as in Ref. 110), and also "climatic optima". He notes almost identical conditions prevailing, for example, in the Ordovician and Triassic, and in the Upper Devonian and Lower Paleogene. He distinguished climatic optima of the first and second kind. The first includes the Triassic and the second the Upper Cretaceous. The first, characterized by the maximum development of zonal circulation, occurs at the time of the lowest solar activity.

In this sense a climatic optimum of the first kind, as defined by Predtechenskiy, coincides with interglacial conditions as defined by Willett. A climatic optimum of the second kind develops under

conditions when there is maximum interlatitudinal exchange, which should coincide with maximum solar activity. Thus, such an epoch corresponds to that which Willett considers "intermediate".

According to Predtechenskiy, the great glaciations occur when there is a singular equilibrium between zonal and meridional processes, that is, when there is moderate solar activity. In this respect the viewpoints of Willett and Predtechenskiy coincide. A weak point in these hypotheses is the neglecting of the results of paleogeography. In actuality, different arrangements of the continental masses and inland seas and a difference from the present-day direction of sea currents in the geological past indubitably were extremely important factors for the climates of the past.

It is well known that Brooks (Ref. 112) held to the hypothesis in his hierarchy of climatic changes that climatic cycles with a duration of several decades, centuries and possibly millennia were associated with variations of solar activity. However, Brooks felt that cycles with a duration of several tens of thousands of years were associated instead with variations of the earth's orbital elements, that is, in this respect he held the same view as expressed by M. Milankovich (Ref. 113). With respect to climatic variations with still longer periods, Brooks felt that it is geological phenomena such as mountain formation which are the principal cause. A different position has been taken by Flohn (Ref. 114) and Viete (Ref. 115). They feel that the great climatic changes of the geological past can be attributed to two factors: (a) according to Brooks, to geological and geotectonic factors, and (b) to extremely long cycles of solar activity.

We will now consider the principal results of study of tropospheric manifestations of the 11-, 22- and 5-6-year cycles during recent years, and also to what extent the well-known Brückner cycle is of solar origin.

Wexler (Refs. 116, 117) formulated the problem of more precise determination of the manifestation of 11-year cycles in meteorological phenomena. Wexler had at his disposal more uniform and extensive data than did Clayton. According to the data for 1900-1939, the difference in mean pressure values at the maxima and minima of the 11-year cycles for both the winter and summer is maximum at 60-70°N. In winter this difference is greater than in summer; at the latitudes mentioned it is ≈ 0.8 mb. At latitudes 40-45° the difference becomes equal to zero, and between parallels 30-35° it attains a maximum negative value (≈ 0.2 mb). A separate study of winter and summer data is a step forward in comparison with Clayton's studies. Wexler obtained a still clearer picture by making a separate study for the months of January. In this case the difference attains +2 mb at latitude 70° and -0.7 mb at latitude 40°. A longitude effect also is very conspicuous. By analogy

with Clayton's work the formation of like differences also was done for temperature and precipitation. The results introduce refinements into Clayton's conclusions. A study of the differences between the conditions for high and low solar activity in July reveals that for this month the effect is considerably less clear. The problem of the stability of 11-year variations of the meteorological elements also has been investigated by Goldschmidt and Reuter (Ref. 118). They analyzed a winter temperature series for central Germany dating back to 1828.

It was found that the 11-year recurrence of cold winters was expressed more or less clearly between 1828 and 1870. Beginning in 1870, several shorter cycles were observed, that is, the 11-year cycles were mixed with 7-year cycles. It should be noted that in the period from 1828 through 1870 there also were such cycles, as well as cycles with durations of five and eight years, but they were not so frequent as the 7-years cycles after 1870. Precipitation in the neighborhood of the Mediterranean Sea also is related to solar activity.

In the investigations which were made by Carapiperis in Greece (Ref. 119) and Godoli in Italy (Ref. 120), it was found that summer precipitation reveals a clearly expressed negative correlation with solar activity, and thus has a variation in the 11-year cycle opposite to the variation of Wolf numbers. On the other hand, winter precipitation has a positive correlation with R, but it is less close. However, a single-peaked curve is by no means always obtained, since in places there is a double or triple wave in the 11-year cycle.

There is a particularly clearly expressed negative correlation between the intensity of summer precipitation and solar activity in regions with a desert climate, such as in Libya. Curious results are obtained from a refinement of the behavior of thunderstorms in the 11-year cycle (in comparison with that which was obtained earlier by Brooks in Ref. 14).

Polli (Ref. 121) investigated the variation of the number of thunderstorms in the 11-year cycle for 12 European stations situated between Leningrad and Naples. His principal conclusion was that in the north the number of thunderstorms follows solar activity, but in the south the pattern is the opposite. At the stations of Central Europe there is a combination of the first and second types; in that area secondary waves develop and the relationship with solar activity is expressed weakly, which can be seen even from the old study made by Myrbach (Ref. 122). Polli believes that in Europe the parallel variation of the number of thunderstorms and solar activity is expressed most clearly at Leningrad and is opposite at Naples. It is interesting to note that A. P. Moiseyev (Ref. 123) also observed parallelism in the frequency of thunderstorms and solar activity, but nevertheless at Moscow secondary waves already can be observed from time-to-time in the 11-year cycle.

L. A. Vitel's has obtained a very clearly expressed curve of the frequency of deep cyclones in the 11-year cycle. The minimum of this frequency occurs in the epoch of the minimum of the $\bar{a}$ index, with the principal maximum being observed on the descending branch of the cycle. A corresponding diagram has been reproduced repeatedly in monographs on solar-tropospheric relationships (Refs. 1, 38). However, in another study (Ref. 124) Vitel's has shown that the depth of cyclones is related nonlinearly to the quantity of precipitation falling in a cyclone.

The variation of the quantity of precipitation in the 11-year cycle can be highly varied. Vitel's believes that there can be 32 types of the 11-year cyclic curve of precipitation, depending on how cyclones and anticyclones behave in the 11-year cycle. Among the other investigations devoted to macrosynoptic characteristics in connection with the 11-year cycle is the work of Z. Gregor and L. Krzhivskiy (Ref. 125), in which they establish a relationship between the frequency of occurrence of fronts over the eastern part of the Atlantic Ocean and over Europe, and solar activity. Strictly speaking, they do not consider the fronts as a whole, but segments of fronts with a latitudinal extent of 5° and a longitudinal extent of 10°. As the characteristic of solar activity the authors used the $\bar{a}$ index. They were able to show that in years of a minimum the fronts lie to the east of the Azores and over the Mediterranean. Another frontal zone is observed over the Norwegian Sea, Scandinavia and Finland. In an epoch of maxima there is only one zone over the North Atlantic. The relationship is closest for January.

Bodurtha (Ref. 126) investigated the characteristics of anticyclones in Alaska and in part of Canada. The map of pressure distribution in these regions for December through February of years with low solar activity differs appreciably from the corresponding map for epochs of high solar activity. Anticyclogenesis is emphasized strongly in Alaska in years of high solar activity and is weakened appreciably in years of low activity. In this investigation Wolf numbers were used as an index of solar activity.

Among the other investigations of manifestations of the 11-year cycle is the work of N. A. Belinskiy and M. G. Glagoleva (Ref. 127). They considered cyclonic and anticyclonic anomalies within the limits of the European natural synoptic region. By averaging the corresponding characteristics month by month these authors then expanded the derived data into series of Chebyshev polynomials, retaining terms to the third degree. A 10-year cycle is found in the coefficients of the polynomials when this method was used for processing data for 1928-1938 and 1941-1958.

A study that belongs to the same group of investigations as the mentioned work of Wexler is that made by Brier (Ref. 128). This was a study of the difference in surface pressure along circles of latitude for the maxima and minima of 11-year cycles. The author used years of

maxima $\pm$ 1 year and years of minima $\pm$ 1 year for the epoch 1899-1939. In the high latitudes this difference was positive, that is, atmospheric pressure in the epoch of a maximum was higher than in an epoch of a minimum and the value of the difference was strongly dependent on the season: in the winter (October-March) it was 1 mb, and in summer (May-September) about 0.5 mb. Between parallels 50° and 20° this difference is negative and is 0.1 mb in winter and a little greater in summer.

With respect to temperature, in the northern part of both Europe and the United States this difference is negative, that is, near the maxima of solar activity the temperature in both regions is lower than near the minima. Dammann, who processed precipitation series for German stations (Refs. 129, 130), demonstrated that there is an 11-year cycle in the variations of summer precipitation. Investigators have given their attention to the tropical zone. Attempts have been made to establish the cause for the worsening of those correlations which were characteristic of Africa at the beginning of the current century. Investigations, therefore, have been made using more recent data, partly using aerological methods.

Sen-Gupta (Ref. 131) has made an investigation of the autumn storms of the Bay of Bengal in epochs of low and high solar activity. Near the minimum of the 11-year cycle the number of storms increases, their trajectories become closed and they give rise to considerable precipitation on the eastern coast of India. Aerological investigations of the regions of formation of these storms have shown that in the epochs of minima of solar activity the temperature at this same height is lower than in epochs of maxima.

Stranz (Ref. 132) cites data on the heights of the tropopause at Leopoldville as revealed by radiosonde data from June, 1953 through December, 1958. He discovered a correlation between this characteristic and sunspot number. In the epoch of the maximum of the 11-year cycle the tropopause is higher and the vertical temperature gradient is lower. The reverse picture is observed at the minimum of the cycle.

Also of a certain interest is the investigation of regional manifestations of solar activity in the troposphere. Whereas in the earlier stages of development of the sun-troposphere problem, when there was a question as to the basic reality of solar-meteorological relationships, such studies were not very convincing, they now have the right to be considered. As an example of such an investigation, made in 1955, we can mention the work of Peisino (Ref. 133), who investigated the mean annual differences between nighttime and daytime precipitation and solar activity for a station in Italy. Using data for 1925-1954, he discovered that this difference attains a maximum at the minima of solar activity.

Among the investigations of the 11-year cycle in the hydrological regime made in recent years is the 11-year variation in the severity of ice conditions in Arctic seas, noted by I. V. Maksimov (Ref. 134), and the conclusion drawn by A. V. Shnitnikov that there is a manifestation of the principal solar cycle in the level of Scandinavian lakes (Ref. 135).

Perhaps the most interesting result of study of manifestations of the solar cycle in the troposphere during the last 10-15 years, excluding the revelation of the role of long-term (80-90-year and longer) solar cycles, has been the establishment of fully convincing conclusions concerning the role of the 22-year cycle in meteorological phenomena. In this connection particular attention should be given to the series of studies made by Willett (Refs. 136, 137). He compiled maps similar to those of Clayton, but for the 22-year cycles, that is, he made a study of those differences which are detected in the fields of the principal meteorological elements at the time of transition from one (such as an even) 11-year cycle to the next (odd).

Although, as we have seen from Chapter 1, the amplitude of the 22-year cycle is in no way comparable to the amplitude of the 11-year cycle, the tropospheric effect of this double cycle is fully commensurable with the effect of the 11-year cycle, and in certain cases (definite regions, seasons, etc.) the effect of the 22-year cycle is considerably greater than that of the 11-year cycle. Willett believes that with a transition from the minimum of an 11-year cycle to a high maximum (that is, in an odd 11-year cycle) there is an increase of active corpuscular solar radiation, leading to an increase of atmospheric pressure in the high latitudes and the development of conditions favorable for meridional circulation, blocking situations, etc.

On the other hand, the ascending branch of even 11-year cycles is characterized by an increase at that time of solar active wave radiation; and this stimulates the development of zonal circulation with its subsequent displacement into increasingly higher geographic latitudes The maps prepared by Willett support the validity of such a model, but in reality the pattern should be more complex.

We will now present certain investigation results on the manifestation of the 22-year cycle in the troposphere that were obtained by other authors. Krjivsky (Refs. 138, 139), who investigated the long-term variation of precipitation in Europe, established the existence of regions in which a 22-23-year cycle, rather than an 11-year cycle, is manifestated. The demarcation line between these two groups of regions passes across Iceland, the central part of Scandinavia, and Germany; to the east of this line there is a 22-year cycle, and to the west an 11-year cycle.

Simojoki (Ref. 140) analyzed the temperatures of Stockholm for the months of December-March from 1756 through 1950. In addition, he analyzed the mean annual levels of the Baltic Sea for the period from 1811 through 1943 on the basis of data for Swinemunde. In both cases he discovered a 22-23-year wave, and the 22-year cycle is expressed more clearly in the variations of the level of the Baltic Sea.

Xanthakis (Ref. 141) considered the following characteristic of annual temperatures: he introduced the values $R = T_7 - T_1$ and $R' = T_6 - T_{12}$, that is, the differences of the mean temperatures for July-January and June-December. He then introduced the values $T_e = 1/2\ (T_3 + T_4)$, that is, the half-sum of March and April temperatures, and $T_e' = 1/2(T_9 + T_{10})$, that is, the half-sum of September and October temperatures. After this he formed the value $X = (T_e' - T_e) - 1/2\ (R - R')$.

The values X were averaged for 13 stations of Europe and the Near East. If X is plotted as a function of the year, and the mean annual Wolf numbers are plotted on this same graph, with a displacement of X by 11 years to the left we obtain a fully satisfactory correlation. Making this same computation for the value X, but for stations in the United States, Xanthakis noted that in this case there is a correlation without a displacement. A direct comparison of X for Europe and the United States reveals a phase difference equal to one 11-year cycle. Xanthakis therefore concludes that for certain characteristics of the annual temperatures of Europe and the Near East, a 22-year solar cycle is characteristic, whereas for the United States the 11-year cycle is characteristic.

In the work of I. M. Soskin, mentioned in connection with the 80-90-year cycle, where the author studied the relationship between solar activity and the hydrological conditions of the Barents Sea, an investigation was also made of the influence of the 11-year cycle on these conditions. It was found that there is an inconstancy with respect to phase: one 11-year cycle is characterized by certain conditions during a specific phase, and the next cycle by different conditions; but after two 11-year cycles the conditions are repeated. A 22-year cycle, therefore, was also detected in this case.

We will now discuss the Brückner cycle and its possible heliophysical nature. M. S. Eygenson (Ref. 143) has written recently in support of such an interpretation of the 35-36-year climatic cycle discovered late in the last century by Brückner (Ref. 142).

It should be noted that the strong warming of the high latitudes in the 1920's and 1930's in all probability was the reason that the dry

period, which should have appeared in the late 1920's in accordance with the Brückner cycle, did not occur. This circumstance, as well as the results of a careful, but somewhat formal investigation by Wagner (Ref. 144), cast doubts on the reality of the Brückner cycle. At the same time, a number of new investigations tended to support the reality of such a cycle.

However, it is obvious that it is impossible to insist on a cycle of exactly 35-36 years. It is necessary to assume, as A. V. Shnitnikov did (Ref. 145), that this cycle covers 40 to 60 years. There is, therefore, no need for identification of the Brückner cycle with hypothetical solar cycles having a duration of about 33 years, which apparently were discovered by Lockyer and others (see Chapter 1).

We feel that it would be more correct to agree with M. S. Eygenson that the Brückner cycle is part of the 80-90-year cycle. From the approximate physical model of climatic variations that we described in the preceding section of this chapter on the basis of the Defant and Angström hypotheses, we find that the single 80-90-year wave of solar activity is entirely capable of giving rise to two, and in certain cases three waves in the long-term change of the fields of meteorological elements.

It must be remembered that in equation (5.3) for interlatitudinal macroturbulent heat exchange the time scale for change of each of the cofactors entering into the right-hand side will be different. The coefficient of horizontal macroturbulent exchange A, to that degree which is caused by processes occurring in the high layers of the atmosphere, i.e., in the long run caused by solar activity, should be more unstable than the meridional temperature gradient. This gradient is to a considerable extent dependent on the underlying surface, i.e., factors having a far greater inertia. This circumstance apparently explains how a double and sometimes a triple climatic wave can develop from a simple solar wave.

The problem of the 44-year cycle discovered by Cullmer in the paths of North American cyclones (Ref. 146) also possibly should be considered from this point of view. However, it is not impossible that in this particular case there is a real 44-year solar cycle of the type noted by A. Ya. Bezrukova. It should be noted that if the above-cited interpretation of the Brückner cycle is correct, this indirectly confirms the "supersecular" stability of the 80-90-year cycle in that cycles very close in duration to the Brückner cycle have been discovered by Lunsgersgauzen (Ref. 147) on the basis of geological data for the Precambrian.

During recent years considerable attention has been given to the discovery of a 5-6-year cycle in various meteorological and

macrosynoptic processes. Attempts also have been made to trace even shorter cycles in atmospheric phenomena that certain investigators feel are of solar origin. Among the authors of such studies we should mention M. S. Eygenson (Ref. 38) in particular. He feels that the 5-6-year cycle is of solar origin and is manifested in the variation of the $\bar{a}$ index. In Eygenson's opinion, this index reflects the activity of solar corpuscular radiation, and the conclusion, therefore, can be drawn that corpuscular radiation itself has a 5-6-year cycle. The first maximum of the $\bar{a}$ index, and this means the first maximum of solar corpuscular radiation, occurs in the epoch of the Wolf number maximum. The second occurs on the descending branch of the 11-year cycle, that is, when the radiator is in the most favorable position relative to the earth.

It should be noted that certain data on magnetic disturbances apparently confirm the point of view (Ref. 148) that magnetic storms with a sudden commencement predominate in the high phases of the 11-year cycle and storms with a gradual commencement, associated with a more directed corpuscular radiation, predominate at the end of the 11-year cycle. However, these conclusions cannot be considered final. Baur also is a supporter of the solar origin of the 5-6-year cycle in meteorological processes.

After an analysis of a 124-year series of observations for 14 German meteorological stations, Baur reached the following conclusion (Ref. 149). Not a single dry year was observed in the limits of 0.2-1.0 year from the extremum during the entire 124 years. Within the 11-year cycle there are frequently two maxima and two minima of the recurrence of dry years. Such a double cycle is observed not only in precipitation in Germany, but also in a number of other phenomena: in the winter values of gradients between Iceland and the Azores, the annual variation of pressure in the subtropics of the southern hemisphere, the frequency of exceptionally severe winters in Central Europe and the northeastern United States, the recurrence of unusually warm months in Central Europe, etc.

In a recently published article (Ref. 150), Baur categorically insists on the absence of an 11-year cycle in meteorological phenomena, pointing out that the amplitudes of the 11-year change of meteorological and circulatory characteristics fall within the limits of error. At the same time, he stubbornly defends the idea of the existence of a double, and in certain cases a triple wave within the 11-year cycle. Baur feels that the most clear example of such behavior of meteorological characteristics within the 11-year cycle is a double wave in the alternation of cold winters in Central Europe. After processing data for 10 stations in Central Europe for the period 1804-1958 Baur even developed certain prognostic rules on the basis of his conclusions concerning double and more complex waves of meteorological characteristics in the 11-year cycle. He concluded that in Central Europe the second spring after the

sunspot maximum is characterized by an increase of precipitation, and the second summer after the maximum is dry.

Among Baur's other conclusions are the following: near the extrema of the solar cycles zonal circulation weakens and monsoon (by which is meant Vangengeym's meridional and easterly circulation types) intensifies. An intensification of zonal (westerly) and an attenuation of monsoon circulation is observed twice and sometimes three times between the extrema of the 11-year cycle. Baur formulates certain prognostic rules (already mentioned), but notes that they are usable only when there is a low level of Wolf numbers: from 10.0 to 45.0 (Refs. 151, 152).

Baur repeatedly proposed a hypothesis of the nature of the double wave of active solar radiation in the 11-year cycle. The essence of this hypothesis is that a cyclic increase of solar short-wave radiation is combined with an increase of absorption in the solar atmosphere. However, from an astrophysical point of view such a hypothesis is improbable (Ref. 153).

Among the studies that discuss the joint manifestation of the 11- and 5-6-year cycles is a paper by T. V. Pokrovskaya (Ref. 154). After breaking down each 11-year cycle into four parts (minimum, ascending branch, maximum and descending branch), Pokrovskaya calculated the recurrence of monthly temperatures for the USSR for each group of individual phases. It was found that there is a double wave in the 11-year cycle. The annual variation of temperature deviations from the norm gives a single wave in the 11-year cycle, although for certain stations the dependence is more complex. The correlation of mean monthly temperatures with the phase of the 11-year cycle, evaluated using a correlation ratio, is in general small (the maximum value is 0.28). The maximum correlation is observed in February-March and the minimum in January, with a secondary minimum in July.

Visser made an attempt to detect cycles shorter than 11 years in certain meteorological phenomena. He investigated a 3-year cycle of high atmospheric pressure and a 7-year recurrence of falling of precipitation in Peru on the basis of data for 1833-1953. When the Wolf number is > 85, the 3-year intervals between years with high pressure are expressed quite clearly, but when the number is < 85 the high pressure is repeated for the most part after two years. Years with high levels of precipitation in Peru for the most part fall in epochs of low solar activity (Ref. 155).

Berkesz, using data for 1780-1959 (Ref. 156), found the presence of a stable 4-year period in the temperatures of Hungary. In the course of 180 years 46 cases were observed when February was colder than January. Berkesz associates these cases with variations in solar activity.

Closely related to the problem of the 5-6-year and shorter cycles in the troposphere is the problem of the relationship of meteorological phenomena to intra-annual variations of solar activity. The author was concerned with this problem as early as 1940 (Ref. 157). At that time it was possible to demonstrate that over a large part of the USSR cold springs are observed considerably more frequently when there is a positive fluctuation of solar activity than when there is a low level of activity at that time of year. There is a particularly clear pattern for May at stations such as Moscow and Leningrad. This problem later was investigated by Baur (Ref. 151).

In his investigations Baur was able to establish that if a sum is formed of the mean monthly Wolf numbers for December, January, February, March, April and May of a particular year (the December is that of the preceding year), and a sum then is obtained for May, June, July, August and September of the same year, if the second sum exceeds the first by more than 4.0 in Wolf numbers, and if at the same time the sum of the mean monthly Wolf number values for August-September exceeds by more than these same 4 units the same sum for October-November, a cold winter can be expected in Central Europe.

This is true provided the following additional conditions prevail: (a) the mean monthly value of the Wolf number for August-September is not greater than 110; (b) the November mean pressure differences between the Azores and Iceland (Ponta Delgada and Stikkisholm) do not have a negative anomaly exceeding 8.5 mb; and, (c) the mean temperature of Berlin for the ten-day period from December 1-December 10 does not exceed the norm by more than 2.5°.

Despite the complexity of this prognostic scheme, it is indisputably of interest to do research in this direction. For this reason the author undertook an investigation in 1959 of the influence of intra-annual fluctuations of solar activity on large-scale meteorological processes (Ref. 158). Specifically, the problem formulated was determination of the influence of the intra-annual variation of solar activity on the type of tropospheric circulation in the northern hemisphere.

The direct purpose of the study was a clarification of the extent to which monthly positive anomalies of each of the four groups of circulatory mechanisms (as classified by Dzerdzeyevskiy) are related to solar activity. This was done by computing the frequency of each of the groups of circulatory mechanisms for 50 years (1899-1948). A mean was determined, the deviations for the corresponding months of each year established, and standard deviations computed. The corresponding months of individual years were compared with these data. If the deviation from the mean for a particular group of circulatory mechanisms of any year exceeded the standard deviation, the corresponding month of such a year was used as a reference month.

The next method used was that of superposing epochs. Six months were counted off in both directions from the corresponding month of the year in which a positive anomaly had been discovered; the mean monthly observed Wolf number was then noted for each of these 13 months. The reference months then were superposed, but before averaging, the maximum Wolf number value in the particular line was assigned the value 100 percent; the values for the remaining months were expressed in percent of the maximum. This precaution is desirable because the averaging includes years with different levels of solar activity.

A study of the results of investigation of this material by the method of superposing of epochs made it possible to plot a series of extrema (maxima and minima) on the corresponding curves. Since the sample sizes here are very small (the largest does not exceed 11), the investigation of the extrema of the curves can be done only by the small sample method. All extrema can be divided into two categories that reveal both a synchronous and asynchronous relationship to the reference months. In the case of asynchronous relationships those extrema which precede a reference month should be of particular interest. Table 32 gives the results of a statistical evaluation of the extrema of the curves resulting from use of the superposing of epochs method.

Thus, there are six extrema meriting confidence. Of these the two in January and August are related synchronously to positive anomalies of the third and second groups of types of atmospheric circulation, and the remaining four precede anomalies of the recurrence of these groups, and also the first group (zonal circulation). The lag in the tropospheric anomaly relative to the solar extremum is from three to six months, which falls within the limits of the lifetime of an active solar center. It is interesting that intra-annual extrema of solar activity, associated with anomalies of atmospheric circulation, are usually minima.

In particular, it is solar activity minima which are related synchronously to positive anomalies of the third and second groups of types of atmospheric circulation in January and August, respectively. We therefore see that in very important macrometeorological processes solar activity can serve both as an activator and attenuator. The quiet sun can also lead to the development of macrosynoptic processes, but it is obvious that they will be of a completely different character than those associated with a high level of solar activity.

Table 32 shows that one of the clearest examples of an asynchronous relationship when there is a low probability of this relationship having a random character is the increased June frequency of the first group of circulatory mechanisms after a positive extremum of solar activity in March.

Table 32

Group of type of circulation	Reference month	Month of extremum	Type of extremum	Probability of random character of extremum	Time in months between solar extremum and circulatory reference	Number of cases
3rd	January	January	Minimum	0.014	0	9
3rd	February	August	Minimum	0.025	0	8
1st	March	November	Minimum	0.050	4	8
2nd	June	March	Maximum	0.014	3	9
2nd	August	August	Minimum	0.035	0	8
2nd	September	March	Maximum	0.025	6	11

It also is of interest to relate to the intra-annual fluctuations of solar activity not only the positive, but also the negative anomalies of particular groups of types of atmospheric circulation in the northern hemisphere, that is, clarify whether there is a repulsion of certain types of atmospheric circulation by solar activity. Such a study also has been made by the author (Ref. 159). The data and the method used were the same as in Ref. 158.

The author determined the various months of those years in which the frequency of a particular group of types of atmospheric circulation was less than the mean long-term frequency for a particular month by more than one standard deviation. The Wolf number for the particular month was noted for such years, and were also noted for each of the preceding six months and each of the following six months. Thereafter, in order to achieve comparability of individual years, which obviously can fall in different phases of an 11-year cycle, the maximum value of the Wolf number from the 13 months was used as the maximum and the others were expressed as a percent of it. Averaging then was carried out.

The curves constructed on the basis of these mean values were selected on the basis of the following criterion: the extremum of the curve of Wolf numbers reduced in this manner either should precede the reference month in relation to the circulatory anomaly, or should coincide with it. This requirement is satisfied by the months and groups of types of circulation cited in Table 33. This table emphasizes the values of the probability of randomness $\leqslant 0.05$, that is, those cases in which it is possible to assume the relationship between a solar fluctuation and an anomaly of atmospheric circulation to be nonrandom. Cases No. 1 (second group of circulatory mechanisms, January) and No. 10 (third group of circulatory mechanisms, October) have a relationship which lies on the threshold of randomness.

As shown by a comparison of this table and a similar table in Ref. 158, the regular character of the relationship between solar fluctuations and negative anomalies of atmospheric circulation is observed as frequently as is the case of a positive anomaly. However, in the case of positive anomalies there is more frequently a synchronous relationship with the intra-annual extremum of solar activity. Asynchronous relationships with Wolf number fluctuations are more characteristic of negative circulatory anomalies.

As a result of this investigation and Ref. 158, it was possible to compile Table 34, which contains sufficiently reliable cases statistically of the relationship of the anomalies of the frequency of groups of types of atmospheric circulation of the northern hemisphere, and fluctuations of solar activity in different months of the year. The table shows the unambiguous character of the relationship between solar activity and anomalies of the third group of circulatory mechanisms in January.

Table 33

Number of point	Group of type of circulation	Month with anomalous circulation	Month with extremum of solar activity	Type of extremum	Probability of randomness	Time in months between solar extremum and tropospheric reference	Number of cases
1	2nd	January	October	Maximum	0.047	3	5
2	3rd	January	January	Minimum	0.012	0	7
3	2nd	February	August	Minimum	0.034	6	10
4	2nd	April	December	Minimum	0.035	4	8
5	3rd	April	January	Minimum	0.088	3	8
6	3rd	May	January	Minimum	0.140	4	8
7	2nd	June	March	Minimum	0.012	3	9
8	1st	August	March	Maximum	0.025	5	10
9	2nd	October	June	Minimum	0.157	4	6
10	3rd	October	August	Minimum	0.053	2	9
11	2nd	November	May	Maximum	0.015	6	9
12	3rd	December	August	Maximum	0.072	4	6

Table 34

Month of intra-annual extremum of solar activity	Type of extremum	Month of circulatory anomaly	Circulation anomaly	Group of types of circulation
January	Minimum	January	Positive	3rd
August	Minimum	February	Positive	3rd
November	Minimum	March	Positive	1st
January	Minimum	April	Positive	2nd
March	Maximum	June	Positive	1st
August	Minimum	August	Positive	2nd
March	Maximum	September	Positive	2nd
October	Maximum	January	Negative	2nd
January	Minimum	January	Negative	3rd
August	Minimum	February	Negative	2nd
January	Minimum	April	Negative	3rd
March	Minimum	June	Negative	1st
March	Maximum	August	Negative	1st
May	Maximum	November	Negative	2nd

Note: For Januaries the relationship is ambiguous

Table 34 shows that when there is low solar activity in January, in that month there can be both positive and negative anomalies of the number of Arctic intrusions in two or more directions simultaneously. At the present time it is impossible to find the reason for this situation without additional investigations. It is entirely probable that this result indicates only the inadequacy of the criterion of the solar fluctuation without taking into account the level of solar activity and certain other conditions similar to those used in the above-mentioned Baur prognostic rules (Ref. 151).

However, the unambiguous result of other cases indicates there is a possibility of drawing certain conclusions concerning forthcoming circulatory anomalies using the criterion of the intra-annual variation of solar activity. In addition to what has been said concerning the probable intensification of the zonal group of types of circulation in June after a strong solar fluctuation in March, it is possible to mention a whole series of almost equally probable relationships. This same March extremum of solar activity leads to a positive anomaly of Arctic intrusions in one direction in September, and a negative anomaly in June. The latter on a background of intensification of zonal circulation in this month makes it possible to use the March solar extremum to predict the approximate macrosynoptic conditions of June. It also leads to a negative anomaly of zonal circulation in September.

The relative Wolf number maximum in May facilitates a negative anomaly of Arctic intrusions in one direction in November; a decreased solar activity in August creates the prerequisites for an increase in the number of Arctic intrusions in two or more directions simultaneously in the following February; the development of solar fluctuations in October decreases the probability of Arctic intrusions in one direction in January, etc.

We will now attempt to summarize certain results of investigations of this aspect of the sun-troposphere problem in recent years. It is necessary to note first that the "classical" direction of solar-meteorological investigations, that is, the study of manifestations of the 11-year cycle in weather and climate, although not retaining a dominant position among such studies in comparison with the position it held in the preceding period of development of the sun-troposphere problem, still remained under active consideration by investigators.

Studies of particular interest, of course, were those involving a comparison between various indices of solar activity, primarily Wolf numbers and indices of the duration and intensity of solar phenomena, and various kinds of macrosynoptic characteristics, rather than individual meteorological elements. Other investigations of unquestionable interest were those refining the old Clayton data on variations of

pressure, temperature and precipitation in the 11-year cycle (Wexler). Study of manifestations of the 22-year cycle were still more important.

With respect to the 5-6-year cycle, the basic problem remained unsolved: is this cycle of solar or terrestrial origin? It should be stated that the theoretically important problem of the manifestation of intra-annual fluctuations of solar activity in the general circulation of the atmosphere, which also is of exceptional practical importance, is only in the initial stage of its development at the present time.

We will now proceed to the last and most complex problem—the mechanisms involved in the relationship between solar activity and meteorological phenomena. During recent years more attention has been devoted to this problem than ever before, which is entirely understandable because this has been facilitated by the study of the upper layers of the earth's atmosphere, considerable development of various kinds of aerological methods, progress in the field of dynamic meteorology, and, finally, the appearance of a great interest in this problem among highly skilled specialists in the field of atmospheric dynamics.

Whereas the earlier investigations in this field now seem somewhat naive, the studies which have been made in recent years in certain respects can stand up under the most searching criticism, both with respect to the data used and the methods and mathematical rigor used in approaching this particular problem.

A study by A. A. Dmitriyev (Ref. 160) must be regarded as the first attempts to make a rigorous approach to this problem. As a point of departure he used the qualitative empirical results presented on Clayton's maps. In particular, Dmitriyev pointed out that with a cyclic increase of solar activity there is an increase of precipitation in those regions where moist—unstable states usually predominate. This behavior of precipitation indicates that with an intensification of solar activity there is an increase of the coefficient of vertical mixing. Another indicator of an increase of this coefficient with an increase of solar activity is the correlation between thunderstorm activity and solar activity.

Dmitriyev does not attempt to explain how an intensification of active solar radiation leads to an increase of vertical turbulent exchange. From this point of view his investigation has something in common with the previously mentioned work of Angström, with the difference that the latter used as a point of departure an increase of horizontal exchange with intensifying solar activity, whereas Dmitriyev considers vertical mixing. The different dependence of temperature, pressure and precipitation on solar activity for the land and sea make it possible to evaluate the role of this astrophysical factor by using the model used by Dmitriyev for study of the problem of monsoon variations of

pressure. The problem is solved for the flat earth's surface. The shore line is considered straight, coinciding with the Y-axis; the X-axis is perpendicular to the shore line and is directed into the continent; the Z-axis is vertical. Along the shore line at sea level there is a temperature discontinuity 2T cos γt (here t is time, γ is the angular velocity of the earth's orbital motion).

The temperature discontinuity has an annual period of variations, and it is postulated that temperature changes are absent along the surface of the continent with increasing distance from the coast. Velocity becomes equal to zero at the underlying surface and when $z = \infty$. Dimensionless variables are introduced: time, coordinates, pressure, velocity components (motion is plane), temperature and dimensionless angular velocity of orbital motion of the earth, in the latter case the scale being the angular velocity of the earth's rotation on its axis. By writing and solving a corresponding system of equations, and representing the solution through the asymptotic form of the Dirichlet integral, Dmitriyev then introduces additional values characterizing soil temperature and its vertical change.

Other parameters introduced include air and soil density, their heat capacity, soil albedo, and the half-amplitude of its insolation, a characteristic of effective soil radiation and the coefficient of thermal conductivity of the soil. Using such variables Dmitriyev obtains a system of differential equations of the heat conductivity equation type with definite boundary conditions. The value of the vertical coefficient of turbulent viscosity, which is of major interest, is included in the scale factor for conversion of temperature to a dimensionless value. After integrating a system of equations of the heat conductivity equation type and introducing boundary conditions, Dmitriyev obtains the following relations:

$$\tau = \tau_0 \cos \nu t. \tag{5.4}$$

Here τ is dimensionless temperature, determined from the relation

$$\tau = \frac{Tg}{T_0 \omega V \sqrt{\frac{k_x}{k_h}}}, \tag{5.5}$$

where g is acceleration of gravity; T_0 is temperature at some initial time; ω is the angular velocity of rotation of the earth on its axis; V is characteristic velocity; k_x is the coefficient of horizontal heat conductivity; k_h is the coefficient of kinematic turbulent viscosity;

$\nu = \frac{\gamma}{\omega}$; and where γ is the angular velocity of the earth's orbital rotation, and $\bar{t} = \omega t$ is dimensionless time, in this case read from the epoch of the solstice. The expression for τ_0 has the form

$$\tau_0 = \frac{(1-A)\,q\sqrt{2}}{T_0 \frac{\omega V}{g} \sqrt{\frac{k_x}{k_h}\,\omega\nu k'\rho'c'}} \times$$

$$\times \frac{\sqrt{\left(1+\sqrt{\frac{k_h}{k'}}\frac{\rho c}{\rho' c'} + \frac{\sqrt{2}\,R}{\sqrt{\omega\nu k'}\,\rho'c'}\right)^2 + \left(1+\sqrt{\frac{k_h}{k'}}\cdot\frac{\rho c}{\rho'c'}\right)^2}}{\left(1+\sqrt{\frac{k_h}{k'}}\frac{\rho c}{\rho'c'} + \frac{R}{\sqrt{2\omega\nu k'}\rho'c'}\right)^2 + \left(\frac{R}{\sqrt{2\omega\nu k'}\,\rho'c'}\right)^2} \tag{5.6}$$

The expression for νt is as follows:

$$\nu\bar{t} = \nu t_s - \operatorname{arctg}\frac{1+\sqrt{\frac{k_h}{k'}}\frac{\rho c}{\rho'c'}}{1+\sqrt{\frac{k_h}{k'}}\frac{\rho c}{\rho'c'} + \frac{2R}{\sqrt{2\omega\nu k'}\rho'c'}}; \tag{5.7}$$

here A is albedo of the soil; q is the half-amplitude of its insolation; ρ' is soil density and ρ air density; c' and c are the corresponding heat capacities; k' is the coefficient of soil thermal conductivity; R is a coefficient taking into account effective radiation of the soil; and, $\bar{t}_s$ is dimensionless time for the epoch of the solstice.

Proceeding on the basis of formulas (5.5) and (5.6), Dmitriyev qualitatively determines the change of amplitudes of temperature with a change of the kinematic coefficient of turbulent viscosity k_h, which is considered directly dependent on solar activity. This is done by inverse conversion to ordinary variables using the following relation:

$$\tilde{T} = \frac{T_0 \omega V}{g}\sqrt{\frac{k_s}{k_h}}\,\tau_0. \tag{5.8}$$

The sign over the T denotes the amplitude of the changes. After differentiating (5.8) for k_h, Dmitriyev finds that for any positive value $\frac{\partial \tilde{T}}{\partial k_h} < 0$. The value $\tilde{T}$ can be represented in the following form when ($k_h = 0$):

$$\tilde{T}=\frac{2(1-A)q}{\sqrt{\omega\nu k'}\rho'c'}\cdot\frac{1}{\left[1+\frac{R\sqrt{2}}{\sqrt{\omega\nu k'}\,\rho'c'}+\left(\frac{R}{\sqrt{\omega\nu k'}\,\rho'c'}\right)^2\right]^{1/2}}. \tag{5.9}$$

At the same time, when $k_h\to\infty$, $\tilde{T}\to 0$, that is, when there is an increase in solar activity, the annual amplitude of the temperature gap between the land and the sea decreases. This corresponds to the observations in Ref. 1, and is the first important conclusion from the Dmitriyev model. The representation of the solution of the initial system of equations in the form of a Dirichlet integral makes it possible to obtain the ratio of the amplitudes of pressure and temperature (we will omit the details of the transformations). This ratio has the form:

$$\frac{\tilde{p}}{\tilde{\tau}}=-\frac{1}{\sqrt{\nu}}, \tag{5.10}$$

where $\tilde{\tilde{p}}$ is the amplitude of dimensionless pressure.

Taking into account this formula, the modulus of the amplitude of pressure will be:

$$|\tilde{p}|=\sqrt{\frac{2k_h}{k'}}\cdot\frac{(1-A)q\rho g}{\gamma T_0\rho'c'}\times$$
$$\times\frac{\left[\left(1+\sqrt{\frac{k_h}{k'}}\frac{\rho c}{\rho'c'}+\frac{R\sqrt{2}}{\sqrt{\omega\rho k'}\,\rho'c'}\right)^2+\left(1+\sqrt{\frac{k_h}{k'}}\frac{\rho c}{\rho'c'}\right)^2\right]^{1/2}}{\left(1+\sqrt{\frac{k_h}{k'}}\frac{\rho c}{\rho'c'}+\frac{R}{\sqrt{2\omega\nu k'}\,\rho'c'}\right)^2+\left(\frac{R}{\sqrt{2\omega\nu k'}\,\rho'c'}\right)^2}. \tag{5.11}$$

The differentiation of this expression for k_h gives $\frac{\partial|\tilde{p}|}{\partial k_h}>0$,

and:

$$\left.\begin{array}{l} p=0\ \text{when}\ k_h=0,\\ \tilde{p}\to\frac{2q(1-A)g}{c\omega\nu T_0}\ \text{when}\ K\to\infty. \end{array}\right\} \tag{5.12}$$

It therefore follows that at the time of low solar activity the annual amplitude of the variation of pressure is small; with an increase of activity it increases to the limit given by expression (5.12). Thus, an increase of solar activity leads to an increase of the annual amplitude of pressure. This to a certain degree explains the mechanism of the law of accentuation of pressure fields. The qualitative physical model is reduced to the following: an increase in vertical mixing leads to a heat loss from the soil surface, but if the radiation balance is negative, it leads to its intensified influx. This decreases the temperature amplitude. The lost heat, entering the upper layers of the atmosphere, leads to a decrease of pressure over the heated regions,

which should lead to an increase of the annual amplitude of pressure at the earth's surface. Dmitriyev believes that in this way it is possible to explain also the heating of the high latitudes at the time of an increase of solar activity; even if the intensity of general circulation remains unchanged, the intensified influx of heat into the upper layers, as a result of its increased loss from the earth's surface, should lead to an intensified transport of heat into the high latitudes. Here, however, a more powerful mechanism is the increase of macroturbulent horizontal exchange.

Dmitriyev also makes an attempt to proceed to quantitative calculations. Without citing these computations (for lack of space), we note that they resulted in the derivation of a formula making possible the computation of the change of pressure when there are changes in the coefficient of vertical mixing; we will limit ourselves here to reporting the final results. With a change of this value by a factor of 2 (from $12.5 \cdot 10^4$ to $25 \cdot 10^4$ $cm^2 \cdot sec^{-1}$), the maximum difference in surface pressure between the sea and land increases by 0.7 mb. This agrees rather closely with empirical data.

In another investigation (Ref. 161) Dmitriyev turns to a problem directly associated with the interaction of the upper and lower layers of the earth's atmosphere. As we have seen, the above-mentioned first attempt at developing a thermodynamic model of the relationship between solar activity and tropospheric phenomena is based on the influence of the active sun on the coefficient of vertical mixing as a given value. However, the most complex part of the problem of the mechanisms of the relationship—the problem of the propagation of some effect, which unquestionably from time to time takes place, moving from the upper to the lower layers of the earth's atmosphere—was not discussed by the investigator.

A partial filling of this gap was undertaken in a study by Dmitriyev whose content we now intend to discuss. This study considers the degree of influence of the difference in velocities at great heights on the velocities in the lower layers of the atmosphere. An increase in velocity at great heights can be caused, for example, by overheating associated with a corpuscular effect (initially there will be an increase of vertical velocity, followed by an increase of horizontal velocity as a result of satisfaction of the continuity equation). The problem is considered by Dmitriyev as a two-layer problem; the zonal relative pressure gradient increases to the tropopause H, and then decreases to zero at a level Q times higher than the tropopause. Both the increase and decrease of the gradient are assumed linear.

After denoting the complex dimensionless velocity at the height QH by $\overline{V}$, Dmitriyev investigates the influence of this value on the

velocity at the tropopause. After writing and solving an appropriate system of equations with definite boundary conditions, and using the applicable value of the coefficient of kinematic viscosity, Dmitriyev concludes that it is possible to use asymptotic formulas. As a result, it was found that a change in velocity, even at a level of double the height of the tropopause, exerts virtually no influence on velocity at the tropopause.

Therefore, there is certainly no influence exerted by wind changes in the upper layers of the atmosphere on the velocity at the tropopause. Only if it is assumed that the coefficient of turbulent viscosity at the tropopause is three orders of magnitude higher than it actually is, will it be possible to obtain a certain influence of changes in the higher layers on the wind velocity at the tropopause (approximately 5 percent of the increase of wind velocity at great heights). This forced Dmitriyev to reject such a model and seek a different hydrodynamic model.

Two possible models have been considered, although it apparently is impossible at present to make such a study as fully as was done for the model developed in Ref. 160, or even for that just considered. The first of these models involves a study of convection processes on a rotating sphere. A qualitative formula for the relative vorticity at a particular level is derived from the vorticity equation. This formula makes it possible to estimate, for example, at what level cyclonic vorticity changes into anticyclonic vorticity, or vice versa. It is obvious that simplifying assumptions are required for making such an estimate; specifically, in this particular case it is necessary to assume that the vertical velocity component is not dependent on height. The formula has the form

$$\zeta_{\mathrm{rel}}(r, z) = \frac{\bar{w}(z)}{\omega(h)} \frac{\rho(z, r)}{[f(\xi)]\,\rho_h} \zeta^{(h)}_{\mathrm{rel}}[f_h(\xi)] - \omega\left[1 - \frac{\rho(z, r)\,w(z)}{\rho_h[f(\xi)]\,w(h)}\right]. \tag{5.13}$$

In (5.13) and (5.14) w is the vertical and u is the radial coordinate; $\bar{w}$ (z) is the mean vertical velocity; w (h) is the same at the height h; ρ is density; $\zeta^{(h)}_{\mathrm{rel}}$ is relative vorticity at the level h; $\boldsymbol{\omega}$ is the angular velocity of the earth's rotation; and ξ is determined from the relation

$$\xi = \int_0^r \frac{\bar{w}}{\bar{u}}\,dz - \int_0^r \frac{\bar{u}}{\bar{w}}\,dz; \quad f(\xi)_{z=h} = z. \tag{5.14}$$

For h we take the initial level at a small height above the earth's surface. Relative vorticity is expressed in fractions of $\boldsymbol{\omega}$. Dmitriyev

considers the following example: assume that at the level h the vorticity is cyclonic and equal to $1/2\,\omega$. It is found, as follows from formula (5.12), that at the level $\rho/\rho_h = 0.5$ relative vorticity becomes zero, and above it becomes anticyclonic. Dmitriyev assumes that this same formula (5.12) can be used for finding the level at which cyclonic circulation is replaced by anticyclonic circulation. We will return to this problem again when we consider the thermohydrodynamic model proposed by L. R. Rakipova.

The last of the models considered by A. A. Dmitriyev in this connection applies to the problem of the change of zonal circulation at the time of variations of solar activity. As will be shown in Chapter 6, variations of zonal circulation have a universal character, being manifested at the time of variations of zonal activity not only in the earth's atmosphere, but also on Jupiter. Dmitriyev points out that such variations of zonal circulation can be explained qualitatively if it is assumed that the temperature variations caused by solar activity can lead to a radial expansion of the atmosphere. This conclusion is of considerable interest.

Actually, when the Dmitriyev study was written nothing yet was known concerning fluctuations, pulsations or "breathing" of the atmosphere—phenomena whose discovery, in our opinion, constitutes one of the most important results of the investigations made during the IGY period. It is now known that such pulsations actually occur at the time of sudden intensifications of solar activity, and, as noted in Chapter 4, at the level lying somewhat above the E layer, amount to several homogeneous atmospheres. However, apparently a still more effective mechanism is slow pulsations associated with the phase of the 11-year cycle, since, as was shown in Chapter 4, in the upper layers of the atmosphere there is a temperature variation associated with the 11-year cycle.

We will now return to the model proposed by Dmitriyev. If the rate of expansion of the gas envelope is represented as

$$u = u_r(r)\,u_t(t), \tag{5.15}$$

it is possible to obtain a dependence of the pressure variations on the variations of radial velocity. In agreement with Dmitriyev we will proceed from the following system of equations in cylindrical coordinates:

a) equations of motion

$$\begin{gathered} \frac{\partial v}{\partial t} + u\frac{\partial v}{\partial z} + \frac{uv}{r} = 0; \\ \frac{\partial u}{\partial t} + u\frac{\partial u}{\partial z} - \frac{v^2}{r} = -g\left(\frac{r_s}{r}\right)^2 - \frac{1}{\rho}\frac{\partial p}{\partial z}\,; \end{gathered} \tag{5.16}$$

b) continuity equation

$$\frac{\partial \ln \rho}{\partial t} + u\frac{\partial \ln \rho}{\partial r} + \frac{\partial (ur)}{r\partial r} = 0; \tag{5.17}$$

c) equation of state

$$p = \rho RT;$$

d) heat flux equations (usual notations)

$$\frac{dQ}{dt} = c_v \frac{dT}{dt} - A\frac{RT}{\rho}\frac{d\rho}{dt}. \tag{5.18}$$

Introducing the characteristics of nonlinear equations

$$\begin{aligned} \xi &= rv, \\ \eta &= \rho ru, \end{aligned} \quad \zeta = \int_{r_s}^{r} \frac{dr}{u_r} - \int_0^t u_t dt$$

and assuming that when $t=0,\ r_0 = \varphi(\xi_0);\ \rho = \rho_0(r),\ v = v_0(r),\ u = u_0(r)$, we obtain for the density, zonal velocity, and pressure the expressions which follow:

$$\rho(r, t) = \frac{\varphi(\zeta)\, u_0[\varphi(\zeta)]\, \rho_0[\varphi(\zeta)]}{ru_r(r)}, \tag{5.19}$$

$$v(r, t) = \frac{1}{2}\varphi(\xi)\, v_0[\varphi(\zeta)], \tag{5.20}$$

$$p - p_\infty = \int_r^\infty \rho\left\{g\left(\frac{r_s}{r}\right)^2 + \frac{\partial u}{\partial t} + u\frac{\partial u}{\partial z} - \frac{v^2}{r}\right\} dr. \tag{5.21}$$

These formulas show that simple sinusoidal variations u, with time cause variations of pressure with a series of harmonics. Therefore, a single wave of solar activity in the course of an 11-year cycle, for example, can lead to double and even greater variations of macrosynoptic conditions.

Dmitriyev made an estimate of changes in the zonal velocity by use of formula (5.20). It was postulated that at the 10 km level the amplitude of the vertical component of velocity is 1 cm/sec, and that at this height there is a maximum of the initial velocity. In such a case, at a height of 20 km the amplitude of the variations of zonal velocity can attain 10 m/sec. At the same time, at levels lower than 10 km the deviations have the opposite sign and a lesser value.

L. R. Rakipova has used a somewhat different approach to the solution of the problem of the relationship between solar activity and tropospheric processes. The basis of the approach is the formulation of the problem of the interrelationship between the upper and lower layers of the earth's atmosphere. By the time of the publication of Rakipova's first study (1951), there already was available rather considerable factual information concerning such an interrelationship

(Ref. 162). As noted by Rakipova, this information applies to the propagation of large vortical tropospheric disturbances (cyclones and anticyclones) to higher layers of the atmosphere.

On the basis of such facts Rakipova has developed a model of alternating cyclonic and anticyclonic circulations lying atop one another. The boundaries between circulations of the one type or another are layers with temperature inversions. Rakipova considers that the heights of these boundaries are 20, 50, 100 and 250 km. Accordingly, above a tropospheric cyclone in the vortical stage of development an anticyclone will develop at the 10-20 km level (Rakipova refers to this level as the level of the "lower stratosphere"); at heights of 20-50 km ("middle stratosphere"), another cyclone will develop; at heights of 50-100 km ("upper stratosphere"), another anticyclone develops; and, finally, at heights of 100-250 km (ionosphere), there is another cyclone.

In the case of a tropospheric anticyclone in the "lower stratosphere", there will be a cyclone, etc. But using more up-to-date terminology it would be necessary to say that above the cyclone in the troposphere there is an anticyclone in the stratosphere, a cyclone in the lower half of the mesosphere (in the mesoincline), an anticyclone in the upper half of the mesosphere (mesodecline) and a cyclone in the thermosphere.

The decisive role in the development of vortical processes in the upper layers is played by cyclones and anticyclones in the lower layers. Nevertheless, to a certain degree the inverse phenomena also can be of importance, that is, the upper layers can affect the lower. As a result of satisfaction of the continuity equation, the ascending currents in each of the atmospheric layers, such as in the ionosphere, should be accompanied by divergence in the upper half of the particular layer and convergence in the lower half. But such convergence inevitably causes ascending currents in the lower-lying layer and a combination of divergence and convergence on both sides of the separating inversion layer. Such is the theoretical model of the possible mechanism of the relationship between the upper and lower layers of the earth's atmosphere.

The following is the mathematical representation of the model. We will consider a circular cyclone in its vortical stage. The cyclone is considered stationary and has a vertical axis; the process itself is considered nonstationary and adiabatic. Under these conditions the equations of motion are written in a cylindrical coordinate system in the following form:

$$\frac{\partial v_r}{\partial t}+v_r\frac{\partial v_r}{\partial r}-\frac{v_\varphi^2}{2}+\frac{w}{H}\frac{\partial v_r}{\partial \zeta}=-\frac{1}{\rho}\frac{\partial p}{\partial r}+lv_\varphi, \tag{5.22}$$

$$\frac{\partial v_\varphi}{\partial t}+v_r\frac{\partial v_\varphi}{\partial r}+\frac{v_r v_\varphi}{r}+\frac{w}{H}\frac{\partial v_\varphi}{\partial \zeta}=-lv_r, \tag{5.23}$$

$$-\frac{1}{\rho H}\frac{\partial p}{\partial \zeta}-g=0. \tag{5.24}$$

The continuity equation has the form

$$\frac{\partial v_r}{\partial r}+\frac{v_r}{r}+\frac{1}{H}\frac{\partial w}{\partial \zeta}=0. \tag{5.25}$$

The heat flux equation is written as follows

$$\frac{\partial \theta}{\partial t}+v_r\frac{\partial \theta}{\partial r}+\frac{w}{H}\frac{\partial \theta}{\partial \zeta}+\Gamma w-\frac{A}{c_p\rho}\left(v_r\frac{\partial p}{\partial r}+\frac{\partial p}{\partial t}\right)=0. \tag{5.26}$$

The equation of state is used in the regular form.

In these expressions, $\zeta = z/H$ is a dimensionless vertical coordinate; H is the height of the tropopause; v_r, v_φ and w are the velocity components for radius, the tangent to the isobar and vertical; θ is temperature deviation from the initial undisturbed value, $\Gamma=\gamma_a+\frac{\partial T}{\partial z}$; γ_a is the adiabatic temperature gradient; T is the temperature of the undisturbed atmosphere; l is Coriolis force, equal to $2\,\omega \sin\varphi$; and, t is time. For integration of system (5.22)-(5.26) it is necessary to have an expression for vertical velocity. Rakipova uses the following form:

$$w=\alpha\zeta(1-\alpha\zeta)f(t).$$

Thus, w is a function of ζ and t, that is, height and time.

In such a form w becomes equal to zero at the earth's surface, and upward assumes positive values, which, however, decrease with height ($\alpha > 1$). At a certain height, under the tropopause, w becomes equal to zero, and then becomes negative, and at the tropopause is negative. With such a law of change of vertical velocity with height the sign of vertical velocity and its value correspond to a model of alternating convergence and divergence on different sides of the inversion surface. The dependence of w on r, that is, on the distance to the center of the cyclone, is ignored by Rakipova. We will not cite a sample of solution of the system of equations. We will give only Rakipova's final formulas for computation of the temperature and pressure fields from their given values at the times t = 0 and t = t_k, at the center of the cyclone r = 0, and at the distance r_k from the center, and also from the initial

and boundary values for the wind field. Rakipova also derived similar formulas for an anticyclone.

Pressure field (cyclone):

$$\frac{gT_0}{\beta}\left[\left(\frac{p}{p_0}\right)^{\frac{R\beta}{g}}-\left(\frac{P}{p_0}\right)^{\frac{R\beta}{g}}\right]=-gH\zeta-\frac{1}{2}\omega^2r^2+\frac{\omega^2+\frac{a^2}{4H^2}}{2}r^2\zeta^4e^{-2F}\times$$
$$\times\left(\frac{1-\alpha\xi}{\zeta}e^F+\alpha\right)^4+\zeta^3e^{-1.5F}\left(\frac{1-\alpha\zeta}{\zeta}e^F+\alpha\right)^3\int q\,dr-\frac{a}{4H}r^2\times$$
$$\times\zeta^4e^{-2F}\left(\frac{1-\alpha\zeta}{\zeta}e^F+\alpha\right)^3\left(\frac{1-\alpha\zeta}{\zeta}e^F-\alpha\right)\frac{df}{dt}_{t=0}+$$
$$+\frac{ar^2}{4H}(1-2\alpha\zeta)\frac{df}{dt}-\frac{a^2r^2}{8H^2}f^2. \qquad (5.27)$$

Here g is acceleration of gravity; T_0 is the temperature at the initial level; β is the real vertical temperature gradient; P is pressure as a function of time at the center of the cyclone and at the initial level, p is the same, but dependent on all three variables, that is, the sought-for value; R is the specific gas constant; H and ζ, as already denoted, are the height of the tropopause and the dimensionless vertical coordinate; a and α are the constant and a slowly changing parameter in formula (5.27)—the expression for the dependence of vertical velocity on height and time; and, r is the horizontal coordinate (distance from the center of the cyclone), $F=\frac{a}{H}\int f(t)\,dt$, $e=2.718\ldots$, $q=\frac{1}{\iota}\left(\frac{\partial p}{\partial r}\right)_{t=0}$ and, finally, f(t) is a function of time through which is expressed the temporal change of vertical velocity. On the basis of the initial assumption this function is arbitrary, but Rakipova finds that it can be determined through P. For q Rakipova uses a simple expression based on the assumption that the horizontal pressure gradient in a cyclone is proportional to r^2 (with the proportionality factor K).

Temperature field (cyclone):

$$0=\frac{A}{c_p}r^2\left\{\pm\frac{2\alpha a}{H}\left[\frac{df}{dt}-\left(\frac{df}{dt}\right)_{t=0}\right]\zeta(1-\alpha\zeta)\ln\left[r\left(1+\right.\right.\right.$$
$$\left.\left.+\sqrt{1-4\alpha\zeta(1-\alpha\zeta)}\right)\right]-\frac{1}{4}\left[\frac{1}{f}\frac{d^2f}{dt^2}-\left(\frac{d^2f}{dt^2}\right)_{t=0}\right]+K\left[\zeta^3e^{-1.5F}\left(\frac{1-\alpha\zeta}{\zeta}e^F+\right.\right.$$
$$\left.\left.+\alpha\right)^3-1\right]r\ln r\pm\frac{K}{a}\left[\zeta^3e^{-1.5F}\left(\frac{1-\alpha\zeta}{\zeta}e^F+\alpha\right)^2\left(\frac{1-\alpha\zeta}{\zeta}e^F-\alpha\right)-\right.$$
$$\left.\left.-(1-2\alpha\zeta)\right]r\ln\left[r+\sqrt{1-4\alpha\zeta(1-\zeta)}\right]\right\}. \qquad (5.28)$$

Here, as in (5.27), A is the thermal equivalent of work; the other notations are the same as in (5.17).

With respect to the wind field, and with respect to the vertical velocity, it already has been said that the radial component is determined from the continuity equation and has the form

$$v_r = -\frac{ar}{2H}(1-2\alpha\zeta)f(t). \tag{5.29}$$

Tangential velocity is expressed as follows:

$$v_\varphi = -\omega r + \frac{1}{r}\sqrt{A' + B' - D'}, \tag{5.30}$$

where

$$A' = \left(\omega^2 + \frac{a^2}{4H^2}\right) r^4\zeta^4 e^{-2F}\left(\frac{1-\alpha\zeta}{\zeta}e^F + \alpha\right)^4,$$

$$B' = q\left(\frac{1-\alpha\zeta}{\zeta}e^F + \alpha\right)^3 r^3\zeta^3 e^{-1.5F},$$

$$D' = \frac{a}{2H} r^4\zeta^4 e^{-2F}\left(\frac{1-\alpha\zeta}{\zeta}e^F + \alpha\right)^3\left(\frac{1-\alpha\zeta}{\zeta}e^F - \alpha\right)\left(\frac{df}{dt}\right)_{t=0}.$$

For an anticyclone the gradient is directed in the opposite direction, and the components of Coriolis force in equations (5.22) and (5.23) also will have the opposite sign. In this case the principal group of formulas assumes the form:

vertical velocity

$$w = -a\zeta(1-\alpha\zeta)f(t), \tag{5.31}$$

radial velocity

$$v_r = \frac{ar}{2H}(1-2\alpha\zeta)f(t); \tag{5.32}$$

tangential velocity

$$v_\varphi = \omega r - \frac{1}{2}\sqrt{A' - B' + D'}, \tag{5.33}$$

pressure field

$$\frac{gT_0}{\beta}\left[\left(\frac{p}{p_0}\right)^{\frac{R\beta}{g}}-\left(\frac{p}{p_0}\right)^{\frac{R\beta}{g}}\right]=-gH\zeta+\frac{1}{2}w^2r^2-\frac{\left(\omega^2+\frac{a^2}{4H^2}\right)}{2}r^2\zeta^4e^{-2F}\times$$
$$\times\left(\frac{1-\alpha\zeta}{\zeta}e^F+\alpha\right)^4+\zeta^3e^{-1.5F}\left(\frac{1-\alpha\zeta}{\zeta}e^F+\alpha\right)^3\int q\,dr+$$
$$+\frac{ar^2}{4H}(1-2\alpha\zeta)\frac{df}{dt}+\frac{a^2r^2}{8H^2}f^2, \tag{5.34}$$

temperature field

$$\Theta=\frac{A}{c_p}r^2\left\{-\frac{a}{H}\left[\frac{df}{dt}-\left(\frac{df}{dt}\right)_{t=0}\right]\zeta(1-\alpha\zeta)\ln\left[r(1+\right.\right.$$
$$\left.+\sqrt{1-4\alpha\zeta(1-\alpha\zeta)})\right]+\frac{1}{4}\left[\frac{1}{f}\frac{d^2f}{dt^2}-\left(\frac{d^2f}{dt^2}\right)_{t=0}\right]-$$
$$-k\left[\zeta^3e^{-1.5F}\left(\frac{1-\alpha\zeta}{\zeta}e^F+\alpha\right)^3-1\right]r\ln r-$$
$$-\frac{k}{\alpha}\left[\zeta^3\left(\frac{1-\alpha\zeta}{\zeta}e^F+\alpha\right)^2\left(\frac{1-\alpha\zeta}{\zeta}e^F-\alpha\right)e^{-1.5F}-(1-2\alpha\zeta)\right]\times$$
$$\left.\times r\ln\left[r(1+\sqrt{1-4\alpha\zeta(1-\alpha\zeta)})\right]\right\}. \tag{5.35}$$

Having such formulas, Rakipova first computed the distribution of wind, pressure and temperature over a cyclone. She considered the times $t = 0$ and $t = t_k$ for two distances r from the center of the cyclone: $r = 0$ and $r = r_k$. Assuming the vertical dimensionless velocity ζ at the earth's surface to be equal to zero, Rakipova obtained the following values: $t = 0$, $r = 0$, $P_0 = 966$ mb, $\theta = 0$, $r = r_k$, $P_1 = 967.35$ mb, $\theta = 0$, $\frac{v_r}{v_\varphi} = 8.5\cdot 10^{-2}$.

For the next time $t = t_k$:

when $r = 0$, $p = 961$ mb, $\theta = 0$,

when $r = r_k$, $p = 962.35$ mb, $\theta = 0.5^\circ$, $\frac{v_r}{v_\varphi} = 1.5\cdot 10^{-1}$.

In these computations it was assumed that the temperature at the earth's surface is $T_0 = 300^\circ$, at the level of the tropopause $T_H = 240^\circ$ ($H = 6$ km), $t_k = 6$ hours, $r_k = 200$ km.

The computations for the troposphere were reduced by Rakipova to a table (Table 35). The first column gives the height, the second the pressure changes at different heights, and the third the pressure values themselves (in the numerator; its value at the initial time and in the denominator; its value at the final time). The fourth column gives the horizontal differences of pressures at the point $r = r_k$ and at the center at different levels at the initial (numerator) and final (denominator) times. The fifth column corresponds to the second column, but here the change of pressure with time is given not for the center of the cyclone, but for a point situated from the center at the distance $r = r_k$. In exactly the same way the sixth column corresponds to the third. The seventh column gives the temperature deviations for the initial and final times (numerator and denominator respectively) for a point at the distance r_k from the center of the cyclone.

In the next layer—the stratosphere over the tropospheric cyclone —there should be an anticyclone. L. R. Rakipova has computed a similar table for pressure and temperature in the 16-19 km layer. Here $\zeta = 0$ is assumed at the tropopause. The results are given in Table 36. The initial values in this table will be: at the tropopause at the time $t = 0$, the radial velocity is 1.67 m/sec, and the tangential velocity is

Table 35

z_{KM}	$r = 0$		$\Delta_r p$	$r = r_k$		0
	$\Delta_t p$	p		$\Delta_t p$	p	
0	—5.00	966.00 / 961.00	+1.35 / +1.35	—5.00	967.35 / 962.35	0 / —0.40
1	—4.64	860.56 / 855.92	+1.16 / +1.88	—3.92	861.72 / 857.80	0 / —3.38
2	—4.24	763.52 / 759.28	+1.13 / +1.14	—4.23	764.65 / 760.42	0 / —3.98
3	—3.56	674.44 / 670.88	+1.00 / +0.18	—4.38	675.44 / 671.06	0 / —3.18
4	—3.56	593.04 / 589.48	+0.85 / +0.03	—4.38	593.89 / 589.51	0 / —2.40
5	—3.11	518.78 / 519.687	+0.81 / —0.47	—4.39	519.59 / 515.20	0 / —0.03
6	—2.93	451.35 / 448.42	+0.74 / —0.64	—4.31	452.09 / 447.78	0 / +4.79

Table 36

z_{KM}	$r=0$		$\Delta_r p$	$r=r_k$		θ
	$\Delta_t p$	p		$\Delta_t p$	p	
6	—2.93	$\frac{451.35}{448.42}$	$\frac{+0.74}{-0.64}$	—4.31	$\frac{452.09}{447.78}$	$\frac{0}{4.79}$
7	—2.86	$\frac{391.77}{388.91}$	$\frac{-0.31}{-1.57}$	—4.12	$\frac{391.46}{387.34}$	$\frac{0}{8.40}$
9	—2.05	$\frac{294.73}{292.68}$	$\frac{-0.13}{-0.56}$	—2.62	$\frac{294.60}{291.92}$	$\frac{0}{13.5}$
12.5	—1.25	$\frac{179.18}{177.93}$	$\frac{-0.08}{0.00}$	—1.17	$\frac{179.10}{177.93}$	$\frac{0}{14.3}$
16	—0.700	$\frac{108.48}{107.78}$	$\frac{-0.015}{+0.260}$	—0.425	$\frac{108.465}{108.040}$	$\frac{0}{11.4}$
18	—0.532	$\frac{81.958}{81.426}$	$\frac{-0.014}{+0.321}$	—0.197	$\frac{81.944}{81.747}$	$\frac{0}{4.79}$
19	—0.499	$\frac{71.092}{70.593}$	$\frac{-0.009}{+0.402}$	—0.078	$\frac{71.083}{71.005}$	$\frac{0}{0.7}$

9.84 m/sec; at the time $t = t_k$, these values will be 2.05 and -1.71 m/sec, respectively. The maximum value v_z is ±5 cm.

Rakipova also computes similar distributions for the higher layers. The following values are obtained for the mesoincline (19-60 km): at the initial time, the radial velocity at the level of the upper inversion (about 19 km) is v_r = -1.95, and the tangential velocity at this same level is $v\varphi$ = 0.14 m/sec; at the time $t = t_k$, they are equal to -240 and 12.34 m/sec, respectively. The maximum value v_z = 15.1 cm/sec (Table 37).

In the case of the mesodecline (60-100 km, according to Rakipova, although this already passes through the mesopause) the initial values will be: t = 0 at the mesopeak; when r = 0 the radial velocity v_r = 3.20 and the tangential velocity is 2.14 m/sec; at the time $t = t_k$ they are 3.92 and -6.90 m/sec, respectively. The maximum value of vertical velocity is v_z = -24 cm/sec (Table 38). For the ionosphere the

Table 37

$z_{км}$	t	$r = 0$		$\Delta_r p$	$r = r_k$		0
		$\Delta_t p$	p		$\Delta_t p$	p	
19	0	—0.499	71.092	—0.009	—0.078	71.083	0
	t_k		70.593	+0.402		71.005	+0.7
30	0	—0.1043	14.86150	+0.002	—0.05640	14.86170	0
	t_k		14.75720	+0.481		14.80530	—12.40
50	0	—0.0069	1.39044	0.0000	—0.01195	1.39044	0
	t_k		1.38354	—0.0050505		1.37849	+0.65
60	0	—0.002522	0.508381	—0.0000	—0.00613	0.50838	0
	t_k		0.505859	0.00360		0.50226	+25.70

initial values are: at the time $t = 0$ the radial component is 5.33 and the tangential component is -0.36 m/sec; at the time $t = t_k$ they are 6.52 and 30.84 m/sec, respectively; the maximum value of vertical velocity is $v_z = +1.34$ m/sec (Table 39).

Rakipova therefore succeeded in giving a quantitative computation of the effect of the interrelationship between the upper and lower layers of the atmosphere, and showing that in the case of the correctness of the proposed model the influence of the lower layers on the upper layers is quite appreciable. It is obvious that here we are interested in the problem of the possible influence of the upper layers on the lower layers.

Rakipova also made such a computation, and it was found that if solar corpuscular or short-wave radiation is capable of imparting an energy of approximately $5 \cdot 10^6$ ergs to one gram of air at the level of the ionosphere (about 100 km), a cyclone already existing in the troposphere will thereby be deepened by 5 mb. This model was refined at a later date. Among the inversions which Rakipova considers normal (tropopause, inversion at a height of about 20 km, associated with the level of maximum ozone concentration, mesopeak and mesopause), that present in the ionosphere, according to the model, in actuality will be a cyclone, but there can be no inversion at heights of about 20 km, at least not permanently. If this is the case there are only three inversion systems: tropopause, mesopeak and mesopause, and, therefore, the tropospheric cyclone also is accompanied by a cyclone at the E layer level—the ionosphere.

Table 38

$z_{км}$	t	$r=0$: $\Delta_t p$	$r=0$: p	$\Delta_r p$	$r=r_k$: $\Delta_t p$	$r=r_k$: p	θ
60	0	−0.002522	0.508381	0.0000	−0.00613	0.508390	0
	t_k		0.505859	−0.0036		0.502260	+25.7
65	0	−0.001525	0.307403	0.0000	−0.003621	0.307402	0
	t_k		0.305878	−0.002097		0.303781	+49.3
85	0	−0.001206	0.187770	0.0000	+0.0000052	0.0187770	0
	t_k		0.0186564	+0.0001258		0.0187822	+49.3
100	0	−0.000016	0.0030461	0.0000	−0.0000325	0.0030461	0
	t_k		0.0030445	−0.0000309		0.0030136	+47.3

Table 39

$z_{км}$	t	$r=0$: $\Delta_t p$	$r=0$: p	$\Delta_r p$	$r=r_k$: $\Delta_t p$	$r=r_k$: p	θ
100	0	−0.000016	0.0030461	0.0000	0.0000325	0.0030461	0
	t_k		0.0030445	−0.000309		0.0030136	−47.32
200	0	$-0.0119 \cdot 10^{-9}$	$3.80893 \cdot 10^{-9}$	0.0000	$0.4526 \cdot 10^{-9}$	$3.80893 \cdot 10^{-9}$	0
	t_k		$3.79703 \cdot 10^{-9}$	$0.05716 \cdot 10^{-9}$		$3.85419 \cdot 10^{-9}$	−47.20
250	0	—	—	—	—	—	0
	t_k						−186.00

Later Rakipova made an attempt to apply this model to the F layer of the ionosphere as well (Ref. 163); this is illustrated in Table 39. To do this it was necessary to assume that at a height of about 300 km there is a relative temperature minimum. At one time arguments were advanced in support of the existence of such a minimum, but now on the basis of data obtained by direct sounding of the upper layers, it is scarcely possible to assert that anything of the sort actually exists (Ref. 164).

Rakipova's investigation is interesting in that the model gives the most complete dependence of changes of pressure in tropospheric cyclones and anticyclones on heating in a particular ionospheric layer (D, E or F).

The most effective mechanism for the deepening of a tropospheric cyclone is the heating of the D and F layers; heating of the E layer has little effect on its filling. These same layers (D and F) are most characteristic for attenuation of a tropospheric anticyclone; with heating of the E layer the anticyclone intensifies, although not very greatly. The appropriate curves are shown in Figures 54 and 55. The Rakipova studies are the most complete and consistent application of the approaches and methods of dynamic meteorology to the sun-troposphere problem yet made.

However, there are a number of objections to the proposed model. Rakipova has been able to answer a number of them convincingly. For example, the objection that it is necessary to take viscosity and the electromagnetic factor into account in the initial equations when they are used for the high layers of the atmosphere.

There also are more serious objections. The most important of these is that a very great quantity of energy is needed to overcome stable stratification in inversions, and this leads to a considerable decrease of the vertical component of velocity. In essence the picture resembles the conditions of forced convection, when the lifting force works against stable stratification. This circumstance should lead inevitably to an extremely strong attenuation of developing perturbations. Inversions therefore act as a buffer or shock absorber which appreciably attenuates the effect of certain atmospheric layers on others. Such an attenuation of a perturbation should be manifested with particular clarity with respect to the effect of the upper layers of the atmosphere on the lower layers. Taking into account this role of inversions, it is scarcely possible to insist on direct energy relationships between the upper and lower layers, and mechanisms of the catalyst type must be considered. In other words, it is necessary to proceed from the development of tropospheric processes at boundaries of instability, when even a very weak external impulse can lead to the development of a process.

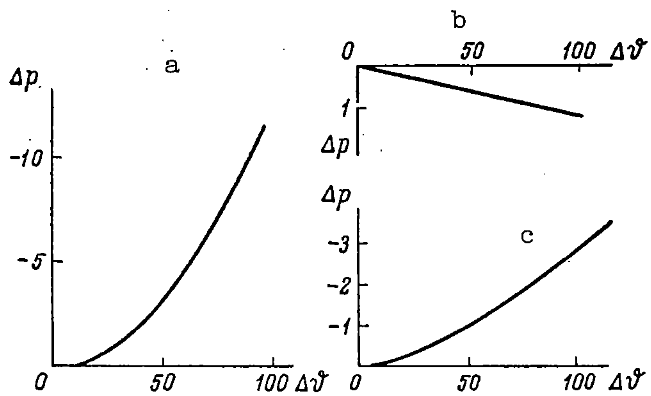


Figure 54. Pressure change in a tropospheric cyclone during heating of layers (according to L. R. Rakipova): a — D, b — E, c — F, in mb

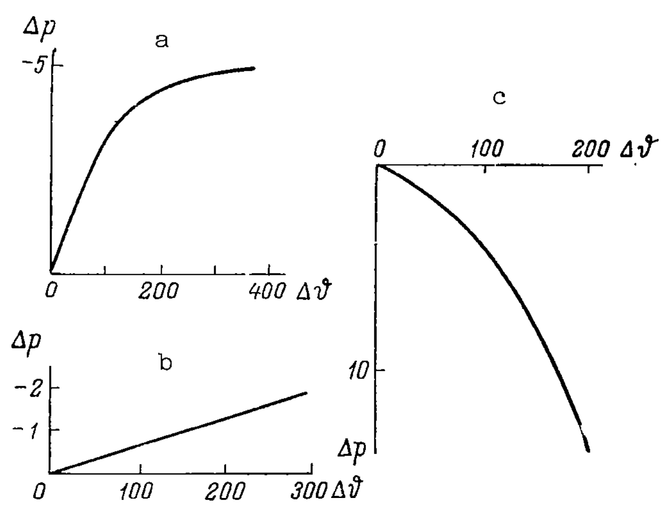


Figure 55. Pressure change in a tropospheric anticyclone during heating of layers (according to L. R. Rakipova): a — D, b — E, c — F, in mb

Such an interpretation, however, leads away from a purely dynamic meteorology solution of the problem, and brings it into the sphere of physical chemistry, that is, in a specific case there no longer is a possibility of more or less reliable quantitative computations. This objection must be considered basic. In addition, although Rakipova cites a number of convincing examples of the interrelationship between the upper and lower layers of the atmosphere, when vortical perturbations are arranged vertical exactly as should be the case according to the basic model (that is, when there is a cyclone in one layer and an anticyclone in the next, or vice versa) it is possible to encounter practical cases frequently when such an alternation of pressure formations is not observed.

We will now investigate another study of this type made by the American meteorologist Spar (Ref. 165). Spar's initial hypotheses are closer to the viewpoint of A. A. Dmitriyev than to Rakipova's. Spar considers a three-layer model of the atmosphere: troposphere, mesoincline and mesodecline. The stratosphere joins with the mesocline. Thus, the lower boundary is the earth's surface, bounded by inversions —tropopause, mesopeak and upper boundary—mesopause. Three models were considered.

In the first model the effect of a thermal perturbation in the mesodecline on circulation is represented. The mesopeak is considered a solid surface; the interaction between the layer in which a thermal perturbation occurred and the underlying layers is not considered.

In the second model it is assumed that there are vertical movements at the mesopeak and tropopause and a vertical propagation of a thermal perturbation, therefore, is possible. It is assumed that the thermal perturbation is situated at the mesopeak (that is, approximately at the level of uppermost atmospheric ozone, approximately at a height of 46 km). The first of these models is baroclinic and the second is barotropic.

The third model is baroclinic; the perturbation is considered the same as in the first model, but in this model vertical movements on the inner boundaries are included. Without discussing the details of the writing and solution of the appropriate systems of equations in Spar's paper, we will note only the final conclusions reached by this author.

1. The absorption of even extremely significant fluxes of energy in the mesodecline is not necessarily accompanied by anomalous heating. This is associated with the compensating effect of adiabatic cooling, jointly with thermal vertical motion.

2. In the three layers of the barotropic model, the thermal perturbation at the mesopeak is attenuated strongly in the direction of the earth's surface. The vertical motions at the two inner boundaries, caused by heating, are very small, insignificant and caused by a tendency to changes of geopotential.

3. The vertical motions on the lower inner boundary in the three-layer baroclinic model are greater than in a barotropic case, but nevertheless are too small to induce an appreciable circulatory process. In this case the attenuation of the perturbations is less appreciable than in a barotropic model. However, for there to be an effect of change of pressure in the lower layers of the earth's atmosphere there is no prerequisite in this case.

On the basis of these considerations, Spar concludes that there is no hydrodynamic mechanism by which the absorption of an anomalous flux of energy in the upper atmosphere can produce any change in atmospheric circulation near the earth's surface, at any rate, until such a time as it is possible to postulate energy fluxes which would be several orders of magnitude greater than those which have been postulated here.

It should be noted that Spar assumes fluxes associated with protons of solar origin that in the mesodecline would have an energy of

10^4 ergs$\cdot$cm$^{-2}\cdot$sec^{-1}. Unfortunately, this figure is almost arbitrary, since for the time being there is no reliable information concerning heating at these heights. The upper part of the mesodecline, adjacent to the mesopause, almost coincides in height with the D region of the ionosphere, but the temperature variations of the latter still have not been studied adequately.

Only recently have data appeared on measurements of the frequency of collisions at these heights, which gives basis for hoping that information soon will be obtained on variations of the temperature regime at these heights associated with solar activity (see Chapter 4).

Ten years ago studies appeared which reported sudden warmings of the stratosphere at the time of intensifications of solar activity. The first example of this kind was the famed "Berlin warming", studied by Scherhag (Ref. 166). There was a temperature increase of 40° per day during the period February 24-25, 1952, at the height lying between the 20 and 10 mb surfaces, i.e., at approximately 27 km. This sudden heating was propagated downward slowly, and after several days led to a temperature increase at the level of the 50-mb surface (20 km) by 12°.

A strong chromospheric flare occurred on the sun on February 24, and was accompanied by a sudden ionospheric disturbance, which was followed later by a considerable increase of geomagnetic activity (an

increase of the Σk_p index by a factor of 3). A careful investigation of circulatory characteristics at the time of this heating, and somewhat later, revealed its considerable influence on intensification of polar easterly winds, the development of deep cyclones over the northwestern part of the Atlantic Ocean, and anticyclogenesis in the Greenland area. Scherhag feels that it is impossible that this warming in the upper layers was of advective origin.

After the "Berlin warming" there were still other examples of such sudden temperature increases at the level of the 55-50 mb surfaces, but they were not expressed so clearly as the case of February 24, 1952. A similar investigation for the Pacific Ocean tropics (neighborhood of the Marshall Islands) was made by Palmer (Ref. 167). He noted that with a sharp intensification of solar activity there is a temperature increase not only above the 55-mb surface, but also between the 300- and 55-mb surfaces. Palmer believes that these temperature increases are of an advective origin and assumes that in the case considered there was a direct influx of heat caused by an intensification of solar short-wave radiation. The influx occurs either directly in the layer in which the heating is noted or in somewhat higher layers.

However, certain authors are skeptical that sudden heatings of layers accessible to aerological observations by ordinary (nonrocket) means can be of direct solar origin. For example, Wexler, who made a special study of the "Berlin warming", does not believe it necessarily was caused by the direct effect of solar activity. Wexler believes that advection can be responsible for this warming (Ref. 168), and that if at the time of a sudden warming of a certain layer of the atmosphere containing not less than 5 percent of the entire atmospheric mass there is an expansion of this layer, and if at the same time the thermal effect can be cumulative, that is, if there is an accumulation of energy in the layer in the course of successive thermal impulses, such a mechanism can be responsible for solar-terrestrial relationships.

Also of considerable interest is a study of this type made in 1959 by Peterson, Jones, Schaefer and Schulte (Ref. 169). These authors made direct observations, using rockets and radiosondes, for investigation of the process of propagation of heating from the upper to the lower layers of the atmosphere. The entire process, observed in late January, 1958, occupied about a week, and was propagated from above downward.

The large time scale of this process is notable. According to the L. R. Rakipova model, the entire phenomenon should occupy perhaps several hours (the duration of the vortical stage of the tropospheric cyclone), or at any rate, not more than a day. The results of the investigation by Peterson and his associates reveal that at heights of 60,

45, and 20 km, the temperature increase is several tens of degrees. In this connection we also can note the studies of Rossi (Ref. 170), Fortak (Ref. 171) and a review of investigations of sharp stratospheric temperature changes made by Vassy (Ref. 172) and others.

The following can be said in summarizing the present-day status of the problem of the relationship between tropospheric phenomena and solar activity. The two most thorough investigations from the standpoint of dynamic meteorology—those by L. R. Rakipova and Spar—stand figuratively poles apart with respect to results and approach to the problem. Whereas the conclusions drawn by Rakipova must be considered too optimistic for the upper atmospheric heat mechanism of solar-tropospheric relationships, the results obtained by Spar appear extremely pessimistic in relation to such a model.

The author feels that the most important requirement for clarification of this extremely important problem is the development of investigations of the type made by Peterson and his colleagues. The data accumulated during the IGY period afford great possibilities in this respect. Only a thorough investigation of the variations of atmospheric temperatures in relationship to intensifications of solar activity at different heights will make possible the solution of the problem of the reality of the thermal mechanism of the relationship between solar activity and tropospheric processes through the upper layers of the atmosphere.

In recent years attempts have been made to consider other possible mechanisms of the relationship between solar activity and meteorological phenomena. We have already noted an attempt to create a new variant of the condensation hypothesis made by Fritz (Ref. 35). In this connection we also should mention an idea advanced by Ney (Ref. 173) that changes in cosmic radiation that actually are observed at the time of strong solar flares can influence the troposphere by means of ionization of the atmosphere. However, this hypothesis is not supported by computations.

In ending this chapter we would like to express certain considerations concerning those directions which we feel are most rational for further development of the sun-troposphere problem. Unquestionably, one of the most important directions is the study of the physical mechanism of the relationship between solar activity and meteorological phenomena. In this work the theoretical investigations should be accompanied by increasing attention to experimental studies, based on vertical soundings of all possible layers of the atmosphere at the time intense processes are occurring on the sun, but also immediately afterwards.

Such an experimental and physicomathematical direction in the sun-troposphere problem, however, by no means should completely supplant empirical-statistical investigations whose success will be largely

dependent on the development and perfection of our concepts concerning both the nature of solar activity, especially its cyclic character, and on the results of purely meteorological and geophysical investigations.

In developing the empirical-statistical direction in the sun-troposphere problem, care must be taken to avoid two extremes: first, "light-weight" studies based on inadequate data which therefore yield unconvincing results, especially if they are not supported by adequate mathematical-statistical criteria for the formulated problem by which the reality of the results can be judged; and second, the use of those statistical methods which may be completely suitable and quite tested for other problems, but which are unsatisfactory when applied to the statistical model that can be involved in the study of solar-meteorological relationships.

Recently many foreign investigators have uncritically, and we might say even dogmatically, used the methods developed by Wiener and his school for the study of random, stationary processes in their investigation of solar-meteorological relationships, especially for determining and clarifying the reality of certain meteorological cycles that have been well established by other methods. In a considerable number of cases these newest statistical methods show that such cycles are not real. But does this correspond to the facts? In a number of cases should it not be assumed that cycles which were determined earlier and whose reality has been verified repeatedly by practical means actually do exist, and the newest statistical methods are not entirely adequate for the formulated problem?

We feel that one of the most promising methods is an investigation of intra-annual fluctuations of solar activity in connection with their manifestation in the lower layers of the earth's atmosphere. In the further development of our concepts concerning the nature of the 11-year solar cycle, it should be the order of the day to study the relationship between solar and meteorological phenomena within the limits of intervals of time commensurable with part of an 11-year cycle.

Also of an unquestionable interest is a further study of manifestations of long-term variations of solar activity in climatic phenomena.

CHAPTER 6
MANIFESTATIONS OF SOLAR ACTIVITY IN THE BRIGHTNESS OF COMETS AND IN PLANETARY ATMOSPHERES

In the two preceding chapters we have discussed certain manifestations of solar activity in the earth's atmosphere. However, at the present time this by no means exhausts all that we know concerning manifestations of solar activity. In the introduction to Chapter 4 it was pointed out that we intended to discuss only one of the large number of problems associated with the effect of solar active radiation on the upper layers of the earth's atmosphere, specifically, the heating of the upper layers. But there is still another important region of manifestations of solar activity—the extraterrestrial.

The planets and their satellites, as well as comets, are affected by active solar influences. This chapter will discuss that problem. Extraterrestrial manifestations of solar activity in many respects represent a completely new field of research. The most valuable information apparently can be expected from measurements made on artificial earth satellites, and especially on space rockets. In this chapter we will discuss only such extraterrestrial manifestations of solar activity as have been studied from the earth.

Section 1. Manifestation of Solar Activity in the Brightness of Comets

That the integral brightness of comets is subject to more or less considerable variations has long been known. However, failures in attempts to discover a correlation with solar activity led Bobrovnikoff (Ref. 1) to the conclusion that the reasons for brightness variations are the result of internal processes in the comet, rather than external effects. The most thorough investigations of this problem during the last 10-15 years have been made by O. V. Dobrovol'skiy (Refs. 2-5). He established that the reasons for Bobrovnikoff's failures, and those of certain other investigators in this field, were associated with a purely empirical approach, and that the specific mechanism of interaction between the sun and cometary matter was not taken into account. We note that from a purely statistical point of view Dobrovol'skiy gave the most rigorous and complete solution of this problem. It is the statistical aspect which we will discuss first.

Proceeding on the initial assumption that the solar agent responsible for variations in the brightness of comets is solar short-wave radiation, Dobrovol'skiy arrives at the following expression relating the flux of this radiation and cometary brightness:

$$m = \text{const} - 5 \lg r - 5 \lg \Delta - 2.5 \lg (S + s), \qquad (6.1)$$

where m is the stellar magnitude of the comet; r is its distance to the sun; Δ is the distance to the earth; S is the area of active regions; and, s is a constant. The area of active regions, as the radiators of active short-wave radiation, can be represented by the area of flocculas, but without great error can be considered equal to the area of faculas.

Thus, the theoretical concepts that we will consider below lead to the conclusion that there is a relationship between the characteristics of solar activity and the stellar magnitude of the comet, and indicate what form it should have. It is clear from (6.1) that this relationship should not be linear, and in actuality, this is found to be true. We will turn first to a comparison of solar activity and cometary brightness over large intervals of time.

Table 40, taken from a study by Dobrovol'skiy (Ref. 2), gives a comparison of solar activity expressed in Wolf numbers and the mean annual number N of discovered comets, together with their mean square deviations. Three gradations of years are considered: with low (from 0 to 10), medium (from 11 to 49) and high (from 50 to 150) solar activity.

This table shows that the difference in the number of discovered comets in years of high and low solar activity is 1.64, almost twice as great as the standard deviation of the difference itself—0.92. Although such an excess of the value of the difference over its standard deviation is not decisive, it nevertheless supports the idea that there is a relationship between the number of discovered comets and the level of solar activity.

Table 40

W	N	σ
0-10	4.36	0.53
11-49	5.50	0.73
50-150	6.00	0.76

Dobrovol'skiy also made a comparison of the brightness curves of individual comets and Wolf numbers. The following points must be remembered. The value S must be taken as it is observed from the comet. Therefore, "the solar activity curve as observed from the earth should be displaced in time relative to the brightness curve of the comet by the value of the difference of the heliographic longitudes of the projections of the earth and comet on the solar disk, divided by the angular velocity of solar rotation" (Ref. 2, p. 27).

When there is an insignificant difference in the heliographic longitudes of the earth and the comet, it can be assumed that the state of solar activity remains the same as when both the earth and comet are projected on the solar meridian. When there is a greater difference in the heliographic longitudes on the solar disk at the time of its "turning" from the direction to the comet to the direction to the earth (or vice versa) there can be rather appreciable changes.

There also is a significance in a difference in heliographic latitudes: at a time when the earth can be projected only on the region of the solar disk falling between $\pm 7^{\circ}$ heliographic latitude, comets as a result of the great inclinations of their orbits can be projected on the high heliographic latitudes. The displacement in heliographic longitude in days is determined using the formula

$$13.4^{\circ}\Delta t = \Omega - A_0 - \text{ar cot } [\tan(\varphi + \omega) \cos i \pm 180], \qquad (6.2)$$

where A_0 is solar longitude; φ is the true anomaly of the comet; Ω, ω,

i are the elements of the cometary orbit (Ω is the longitude of the ascending node, ω is perihelion distance from the node, i is inclination). The inclination of the solar equator for simplicity has been assumed here to equal zero, which of course does not correspond to fact (as is well known, this angle is $\pm 7^{\circ}$). However, this is not reflected in the result of the computations. The heliographic latitude of the comet is given by the formula

$$b = \arcsin [\sin(\varphi + \omega) \sin i] \qquad (6.3)$$

where the notations are the same as for (6.2).

Figure 56 shows a comparison of the variations of solar activity on the basis of Wolf numbers and the brightness of the comet 1936a (Peltier). The interval of observations covers about two months (from June 10 through August 20, 1936). A comparison of the curves shows that an increase in Wolf numbers is accompanied by an increase of cometary brightness. A more careful analysis made it possible to establish that

a good correspondence in this sense is observed when the displacement of the curves of cometary brightness and solar activity do not exceed one or two days.

It should be noted that cometary activity can precede spot-forming activity by an entire solar rotation, that is, cometary brightness increases when a comet is opposite those sectors of the solar photosphere where spot-forming activity does not develop until the next solar rotation. This occurs because in many cases the development of spot-forming activity is prededed (again, by a full solar rotation) by the presence of a facula field in the photosphere or by a floccular field in the chromosphere.

Dobrovol'skiy has analyzed this comet with greater care in Ref. 3. In addition to corrections for the angle earth-sun-comet, all dates of outbursts of cometary brightness were shifted backward by 0.5 day, a value corresponding to the possible time of propagation of active solar radiation. As a result, it was possible to improve further the coincidence of the Wolf number maxima and cometary brightness outbursts.

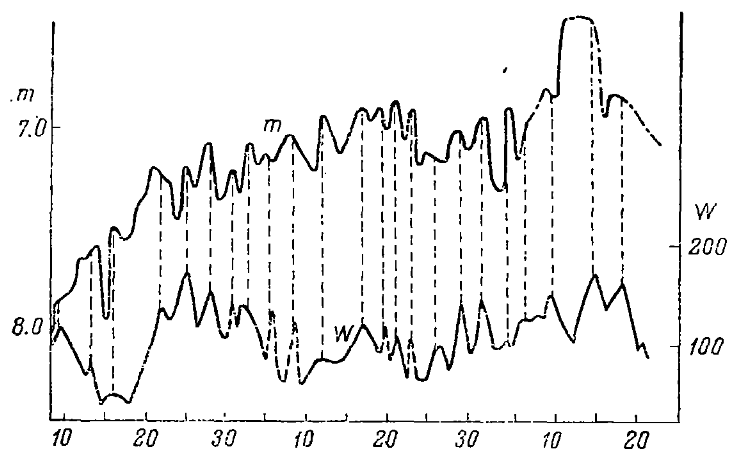


Figure 56. Dependence of the brightness of the comet 1936a (Peltier) on solar activity (according to O. V. Drobrovol'skiy). Left scale—stellar magnitude; right scale—Wolf numbers. The vertical lines correspond to maxima appearing for the first time after one solar rotation

Dobrovol'skiy uses the following method for a rigorous statistical comparison of intensifications of solar and cometary indices. Assuming event A to be an increase of Wolf numbers and event B to be a cometary brightness outburst, it is possible to determine a priori the probability p of their random coincidence as

$$p = 1-q = \frac{T^*}{T} . \tag{6.4}$$

In this process it is necessary to agree about those time gradations within which the events can be regarded as coinciding. Dobrovol'skiy uses one day as such limits, that is, a coincidence is assumed to take place if the events A and B occurred on the same day or if one of them occurred a day earlier or later than the other.

However, if two days or more have passed between events A and B, there is no coincidence. In formula (6.4) T^* denotes the number of days when event A was observed and T denotes the number of days on which there were continuous observations both of cometary brightness and solar activity. In this particular case (the Peltier comet) p = 0.637. Denoting by n the number of events A or B (an interval of time is taken within which these numbers are equal), and by m the number of coincidences within the interval, Dobrovol'skiy determines the correlation coefficient using the expression

$$K = \frac{\frac{m}{n} - p}{\frac{m}{n} + p - 2\frac{m}{n}p}. \tag{6.5}$$

In this particular case K = 0.84. In order to judge the reality of such a correlation coefficient, that is, its nonrandom deviation from 0, it is possible to compute the following criterion:

$$D = \frac{\frac{m}{n} - p}{\sqrt{pq}}. \tag{6.6}$$

In order for the probability of the randomness of deviation of K from 0 to be less than 0.003, it is necessary that $D \geq 3$, since this criterion is based on the 3σ rule. In this case D = 3.0, so that the relationship between brightness outbursts of this comet and Wolf number maxima can be considered statistically established.

A similar analysis was made by Dobrovol'skiy for other comets: 1932V, 1937V (Finsler's comet), 1943 I (Whipple-Fedtke-Tevzadze comet) and others. In addition to comparisons with such solar activity indices as facula area, and especially Wolf numbers, it is of great interest to

seek relationships between cometary brightness intensifications and solar flares. However, the study presented by Houziau and Battiau (Ref. 6) yielded no definite results.

It should be noted that these investigations were formulated in such a way that the solar agent responsible for cometary brightness variations was assumed to be corpuscular streams associated with flares, that is, streams propagating at a rather great velocity (of the order of 10^8 cm/sec) within considerable solid angles and apparently associated with magnetic storms with a sudden commencement and without a 27-day period (corpuscular streams of the first kind). As noted by Dobrovol'skiy, the problem of the presence or absence of a relationship between chromospheric flares and cometary brightness outbursts, that is, between increases in cometary brightness and type-I corpuscular streams, still remains unsolved.

It is of interest to compare intensifications of cometary brightness and geomagnetic activity. Richter (Ref. 8) has investigated brightness outbursts of the Schwassmann-Wachmann comet. During the period from 1940 through 1949 it was possible to establish the dates of onset of 8 such outbursts. Assuming that the velocity of cosmic rays falls in the range 500-1500 m/sec, Richter demonstrated that in all of these eight cases a geomagnetic disturbance can be compared with each brightness outburst. The most probable velocity of corpuscles in this case was 730 km/sec. It should be noted that certain of these geomagnetic disturbances revealed a 27-day period. Thus, the results of an investigation of the Schwassmann-Wachmann comet suggest the presence of a relationship between cometary brightness outbursts and relatively slow corpuscular streams, for which, at the same time, a 27-day period is characteristic.

Dobrovol'skiy (Ref. 5) has compared variations in the brightness of the Finsler and Peltier comets, which he investigated earlier, and geomagnetic disturbances, and arrived at results which in general confirm Richter's point of view. This and other investigations suggest that there is a relationship between intensifications of cometary brightness and type-2 corpuscular streams (that is, streams propagating with a low velocity, revealing a 27-day period, propagating within a relatively small solid angle, and for the most part inducing magnetic storms with a gradual commencement on the earth).

An important aspect of the study of the relationship between cometary brightness and solar activity is an investigation of cometary brightness variations in relation to the phase of the 11-year solar cycle. The Czechoslovakian investigators Link, Vanyssek and Sekanina (Refs. 7, 9, 10) have worked in this direction. But the most complete

investigation from a statistical point of view was made by Dobrovol'skiy (Ref. 4). He first used a large catalog of comets visible with the naked eye during the hundred years from 1850 through 1949.

Dobrovol'skiy's statistical processing of these data is of considerable methodological interest. During the series of years included in the catalog (four years were excluded due to incomplete information concerning them) the total number of cometary appearances was $N_0 = 104$.

All of these appearances were broken down into groups in accordance with the phase of the 11-year cycle. The year of the minimum was used as the zero phase. For a year with the number T the phase was computed using the formula

$$\Phi = \frac{T - T_1}{T_2 - T_1}, \tag{6.7}$$

where T_1 and T_2 are the years of the minimum of the particular and subsequent cycle, respectively. The determined phases were broken down into groups for each 0.1 phase interval, with exception of one group, where it was necessary to use an interval of 0.15. By determining the mean phase in each interval, Dobrovol'skiy then found the mean annual number of comets in each of the phase intervals.

The table that Dobrovol'skiy compiled also gives the absolute number of comets in each of the phase intervals. The result is illustrated in Figure 57, which shows that the number of comets visible to the naked eye as a function of the phase of the cycle reveals two maxima within the limits of a cycle. One of these falls in the phase 0.35, that is, in an epoch close to the maximum of the cycle; the other, less clearly expressed, falls at the end of the cycle.

In order to evaluate the reliability of the result, Dobrovol'skiy uses the variation analysis method. The first step is to find the dispersion by phases of the cycle; this is done by computing the sum of the squares of the deviations of the observed annual numbers from the corresponding means for each phase. In this case the sum was 87.1. Since all the data were broken down into 10 groups, the number of degrees of freedom is 10 - 1 = 9. The dispersion by phases of the cycle is obtained as the quotient from division of 87.1 by 9, which gives 9.68. This dispersion must be compared with the remainder; this is done by dividing the first derived number $N_0 = 104$ by the number of years used, that is, by 96, which gives 1.08.

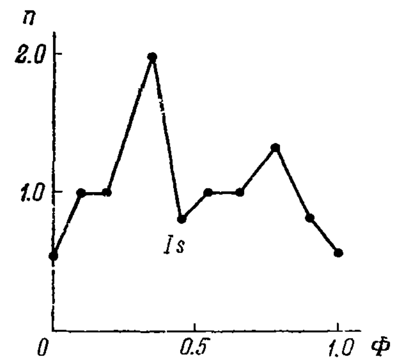


Figure 57. Dependence of the sighting of comets on solar activity (according to O. V. Dobrovol'skiy). Along the x-axis—phase of the 11-year solar cycle; along the y-axis—number of comets visible with the naked eye; s—value of the dispersion assuming a random appearance of comets

The sum of the squares of the deviations of the annual numbers of comets from this value is 101.3. The remaining sum of the squares of the deviations is obtained as the difference of their new number, equal to 96 - 1, and the earlier computed 9, that is, 86. The residual dispersion will be $\frac{14.2}{86} = 0.165$. In order to compare it with the dispersion by phases it is necessary to find the ratio of these dispersions; it is equal to 58.8:1. The phase of the 11-year cycle, therefore, appreciably influences the number of discovered comets, since with such a number of degrees of freedom, if the particular ratio is greater than 2.65:1, the probability of the randomness of the relationship between the phase of the cycle and the number of discovered comets would be less than 0.01.

It also is possible to use a different criterion, that is, find the value

$$\frac{s_\varphi - s_e}{\sqrt{\frac{s_\varphi^2}{2p} + \frac{s_e^2}{2(N_0 - p)}}}, \qquad (6.8)$$

where p is the number of phase groups.

In this case it is equal to 3.9, which is greater than 3, that is, if a normal distribution is not postulated, then the 3σ criterion is satisfied, which indicates a substantial value of the phase of the 11-year cycle for this more general case as well.

The later investigations of Dobrovol'skiy, based on an analysis of a series of cometary observations made by Beyer, confirmed this result (Ref. 5). At the same time, Dobrovol'skiy notes that the dependence of the number of comets on solar activity still has not been clarified in full. All these empirical investigations also are of interest because they were based on a definite physical model of the effect of solar activity on cometary phenomena.

As we have noted, Dobrovol'skiy initially used photodissociation as a point of departure, emphasizing also the excitation of the electron shells of cometary molecules by hard wave radiation. This led him to the derivation of expression (6.1). Thereafter, however, it became increasingly obvious that it was corpuscular radiation which should play the decisive role. The latest discoveries with respect to L_{y_α} and the behavior of its intensity in the 11-year solar cycle indicate in fact that the decisive means by which the sun acts on cometary processes is corpuscular radiation.

At the same time, as noted by Dobrovol'skiy, ultraviolet radiation should be extremely important in the physical chemistry of a cometary head. He was able to show simultaneously that shortest-wave region of the studied solar spectrum, that is, the region of X-radiation, cannot exert an appreciable influence on the change of cometary brightness. However, it is in just this region that there are the most significant variations caused by solar activity.

In the region of waves longer than those of the Lyman series, that is, in the region from 1,300 to 4,000 Å, there are virtually no variations of solar radiation which could be related to its activity, nor are there in the region of still longer wavelengths. However, as pointed out by Dobrovol'skiy, in the case of radiation with a wavelength of > 4,000 Å, the energy of a quantum is too small to cause photochemical reactions (Ref. 2). Thus, the problem of the role of ultraviolet Lyman radiation in the variations of cometary brightness should be clarified finally after the character of the change of this radiation within the limits of the 11-year cycle has been established.

The theory of the interaction of solar corpuscles with cometary matter was developed by Biermann (Ref. 11). He was able to establish that in the case of comets with type-I tails, that is, consisting of CO^+ ions, and to a certain degree N_2^+ and CO_2^+ as well, the solar plasma

interacts with cometary ions, and as a result the latter are entrained by the plasma, and this leads to acceleration directed along the corpuscular stream. Biermann checked this hypothesis using a number of comets: Morehouse (1908 II), Halley (1910 II) and Whipple-Fedtke-Tevzadze (1943 I).

Biermann established that in order to obtain an acceleration of 10^3 g (where g is the acceleration of solar attraction; acceleration of this order of magnitude is typical of type-I tails), it is necessary to have a stream of solar protons with a concentration of 10^4 particles cm^3 at a velocity of 10^8 cm/sec. We note that the value g at a distance of 1 atmospheric unit (a.u.) is 0.59 cm/sec^2. According to Biermann, not only the motion, but even the formation of cometary ions is caused by solar corpuscular streams; the most effective mechanism in this case is charge exchange with a proton of the corpuscular stream.

Whereas numerous comparisons of variations of cometary brightness and solar activity have been made, far fewer studies have been devoted to the relationship between changes in the spectral characteristics of a comet and solar activity. This can be attributed for the most part to the fact that cometary spectra are not too frequently systematically recorded. Nevertheless, certain results have been obtained. It has been found that at the time of brightness outbursts there is a change in the cometary spectrum so that the spectral characteristics of a comet differ somewhat before and after the outburst.

Vogel (Ref. 12) was the first to point out the different character of the spectrum before and after a brightness outburst; this was the result of study of the Pons-Brooks comet. A similar phenomenon also has been observed in other comets (Holmes, Halley). Bobrovnikoff (Ref. 13), who made a special study of the change in the spectrum in Halley's comet, noted an outburst of cometary brightness occurring on May 24, 1910. In this outburst there was a noticeable weakening in the violet and ultraviolet parts of the spectrum and the CN band, and also C_3. It is true, according to Bobrovnikoff, that the outburst of Halley's comet on May 24, 1910 was caused by an explosion in the nucleus of the comet, and it therefore cannot be asserted that in this specific case solar activity was responsible for the change in the cometary spectrum.

Shul'man (Ref. 14) made a spectrophotometric investigation of the Pons-Winnecke comet of 1939. Investigating data for the period from June 11 through June 23, 1939 (during this period the distance of the comet from the sun remained virtually unchanged), Shul'man discovered considerable variations of the intensity of the CN, C_2 and C_3 bands.

It was found that about June 15 or 16 the short-wave part of the spectrum of this comet was attenuated relative to the longer-wave part of the spectrum. It should be noted that in June, 1939, solar activity was low, but on the other hand there was a powerful fluctuation in July, and as a result, in particular, a strong active center was formed which passed across the sun's central meridian between July 7 and 9 (for the terrestrial observer). Consequently, in June this sector of the sun passed across the central meridian between the 10th and 12th.

Taking into account Dobrovol'skiy's assertion that cometary activity can outpace solar activity by an entire solar rotation, it can be considered entirely probable that the changes in the spectral characteristics of the Pons-Winnecke comet, occurring about June 15, 1939, were caused by the mentioned active center, which developed on the next solar rotation, or some other center, appearing in July, 1939, at a time, as already mentioned, when solar activity had intensified appreciably. It is possible to point out a still clearer example of the relationship between the spectral characteristics of a comet and solar activity: I. D. Kupo and V. G. Teyfel' discovered a change in the relative intensity of the band λ 5164 C_2 in the comet 1956h that occurred simultaneously with a change of Wolf numbers (Ref. 15).

In conclusion, we will note still another fact pointed out by Dobrovol'skiy. The velocity of ejection of matter from a comet also is associated with solar activity. If the ejection is caused by corpuscular streams, the velocity should be dependent on the distance of the comet from the sun and the form of the dependence should be:

$$v \sim r^{-\frac{2}{3}}, \tag{6.9}$$

where v is the velocity of ejection, r is the heliocentric distance of the comet. With the fully rational assumption that the concentration of solar protons at the distance of the comet is 10^3 particles/cm^3, the observations fully confirm the law given by formula (6.9).

In summary it can therefore be noted that at the present time the relationship between nonstationary processes in comets and solar activity is established. This important conclusion is evidence that the manifestation of solar activity is universal for the solar system. At the same time, the study of the relationship between solar activity and phenomena associated with comets is one of the means for detecting solar corpuscular streams, which also is of extremely great importance.

Further evidence that the sun and its changing activity constitute a singular "director" of a wide variety of processes in the solar system is the manifestation of solar activity in the atmospheres of the major planets.

Section 2. Manifestation of Solar Activity in Certain Physical Characteristics of the Major Planets

The problem of the relationship between solar activity and phenomena observed on the major planets has been discussed in the literature rather frequently in recent years. But in the past, approximately to the end of the 1940's, investigations of this type were infrequent. Almost all these few studies were devoted to comparisons of solar activity with phenomena on the planets within the 11-year solar cycle. Almost no investigations were made of the manifestation of solar activity within intervals of time commensurable with the cycle of solar rotation. We note in passing that the synodic period of solar rotation for the various planets of the solar system should be different, and differ more or less from the value 27.3 days, representing the mean length of the synodic rotation of the sun for the earth.

The closer the planet is to the sun, the longer should be the synodic period of rotation of the sun for an "observer" located on a planet. For example, it is about 34 days for Mercury and almost 29 days for Venus. On the other hand, for the planets situated at great distances from the sun, the synodic period of solar rotation is close to the sidereal period, that is, approximately 25 days. Almost no individual comparisons have been made of solar and planetary phenomena, which have been found, as we have seen from the preceding section, to be quite successful with respect to solar activity and cometary phenomena.

Until recently there have been no attempts to discover an 80-90-year solar cycle in planetary atmospheres. One of the principal reasons why this exceedingly important and interesting field of the manifestation of solar activity could not be studied successfully was that more or less prolonged and at the same time homogeneous series of planetary observations were (and still remain) very limited.

Solar activity obviously is manifested both on planets having an atmosphere and on those which have none. The recently discovered phenomenon of luminescence on the moon, and also on one of the large satellites of Jupiter, may well be associated with an intensification of solar activity (Refs. 16, 17).

Nevertheless, the manifestation of solar activity on planets possessing an atmosphere should be more varied. This assertion is scarcely in need of additional proof if we recall how many possibilities of the

manifestation of the influences of active solar radiation can be associated with general circulation in a planetary atmosphere (such as on the earth, see preceding chapter). We should therefore briefly mention current concepts concerning planetary atmospheres.

The information currently available to researchers on the Venusian atmosphere is rather modest and contradictory. The assertion that the Venusian surface is always shrouded in clouds became trivial long ago. Venus is always covered with clouds. The question, therefore, of the nature of these clouds naturally arises at once. The answer is dependent directly on which model of the atmosphere is assumed. The latest data, both radioastronomical and that obtained by the American Mariner 2 spacecraft, support the idea that there is a strongly heated surface and a cold atmosphere (Ref. 18). It is therefore even possible that the clouds consist of ice crystals (certainly not water).

The observational data, which for the time being are still difficult to interpret, indicate a diffuse character of the clouds and their considerable variability. It is interesting that the details are detected primarily in ultraviolet light; in yellow light the structure of the cloud cover is scarcely noticeable, and in red and infrared light the structure disappears altogether (Ref. 19).

It is noted that certain hints of a zonal pattern, appearing from time to time on the planetary disk, usually disappear rapidly, which is evidence of a small value of the deflecting force, that is, an indication of slow rotation of the planet (Ref. 19). However, there are still more direct indications on this point; it can now be considered virtually established that the period of rotation of Venus on its axis is long and it it is not equal to the period of its revolution around the sun, i.e., 225 days, in any case it is not less than a week (Ref. 19).

We have already mentioned that current data on the temperature of Venus shows it to be about 300°C (Ref. 18); we note once again that this should be the temperature of the surface of the planet, not its atmosphere. However, at the upper cloud layer level the temperature is 243°K (lower limit) (Ref. 19). Atmospheric pressure near the surface is 20-100 atm. Both radioastronomical data from terrestrial observations and data from the American spacecraft indicate that there are no Venusian radiation belts and no appreciable magnetic field.

With respect to chemical composition, most authors assume that carbon dioxide is a significant component of the Venusian atmosphere; above the visible cloud layer its abundance is from 100 to 1,000 atm-m. We recall that 1 atm-m is the thickness, expressed in meters, of a gas layer which would be obtained if all the particular gas was removed from the atmosphere and compressed to a pressure of 1 atm at a normal temperature.

According to certain opinions, the volume of CO_2 is about 20 percent and the remainder consists of molecular nitrogen (Ref. 19). This component of the Venusian atmosphere will be discussed below in connection with the possible luminescence of nitrogen under the influence of solar activity. There are indications of presence of water vapor and also molecular oxygen (Refs. 20, 19).

At present judgments concerning the Venusian atmospheric circulation are almost entirely speculative. It is true that slow rotation, which can be considered a fact, gives definite indications of the order of magnitude of the deflecting force; the latter will be far weaker than on the earth, which even can lead to direct air exchange between the high and low latitudes. However, recent data on the temperature of the Venusian surface complicate the solution of the problem of circulation, since it becomes unclear where the source of heating is situated and how it is related to the subsolar point.

Proceeding to the Martian atmosphere, by analogy with Venus we should discuss the problem of cloud formations. In the first approximation it can be said that the Martian atmosphere is transparent to the same degree to which the Venusian atmosphere is opaque, as a result of its cloud cover. However, cloud formations do appear from time to time in the Martian atmosphere. On the basis of recent data it is necessary to distinguish three types of cloud formations in the Martian atmosphere.

1. Yellow clouds (assumed to be dust clouds). They can be observed as the winds move them across the planetary disk. These clouds sometimes cover the entire disk of the planet, as was observed for some weeks at the time of the opposition of 1956. It should be noted that dust storms have been observed at the time of earlier oppositions of Mars (Ref. 21).

2. White clouds. These clouds are formed for the most part over the dark regions of the planet. Certain specialists believe that these actually are thin, high clouds, similar to cirrus and cirrostratus clouds on earth. The most acceptable opinion is that they consist of ice crystals; their polarization properties support this hypothesis.

3. Blue haze. Morphologically this is a diffuse substance, with the result that details of the surface disappear in blue and violet light. A blue haze is a far more frequently observed phenomenon than the two preceding cloud forms and it can be said that it is usually present. "Blue clearings" appear only from time to time. According to certain data, the blue haze in the form of diffuse clouds of little contrast and covering great areas is frequently observed near the terminator; it corresponds to sunrise. For now the nature of the blue haze is not understood. It has been postulated that it is related to solar corpuscular streams (Ref. 22). Although there have been objections to this

hypothesis, nevertheless, from this point of view the blue haze can be considered a phenomenon whose comparison with intensifications of solar activity is rather promising.

It is rational to accept a value of 85 mb as the atmospheric pressure prevailing near the Martian surface; this value has been determined from both polarization and photometric data (Refs. 23, 24). This pressure corresponds to that observed in the earth's atmosphere at a height of about 18 km. The mean daily temperature is about $240^{o}K$. At midday in summer in the low latitudes it can attain $300^{o}K$.

No more is known about the chemical composition of the Martian atmosphere than about that of Venus. It has been definitely established that carbon dioxide is present, but whereas for Venus there is a basis for assuming that it is one of the principal components, the situation on Mars is different: the CO_2 content by volume is only 2.2 percent.

Water vapor is present in the Martian atmosphere since it is almost beyond dispute that the polar caps consist of some form of frozen water (ice or frost). Molecular oxygen is also apparently present in some form. However, for the most part the Martian atmosphere most likely consists of molecular nitrogen with a small admixture of argon (Ref. 19).

Circulation of the Martian atmosphere will be discussed below; we note only that the sun is the source of heat for both Mars and the earth, and the rate of axial rotation is almost the same as for the earth, that is, the deflecting force has the same order of magnitude. There is no basis for assuming that the general circulation on Mars will be essentially different from that on the earth.

Planets of the Jupiter type are of particular interest for the problem discussed in this section because it is these, and Jupiter specifically, that reveal the most definite relationships to solar activity. Continuing our scheme of presentation, we will discuss primarily the cloud problem. The cloud systems for planets of the Jupiter group are of particular interest because they are the structural forms observed on these planets. The cloud systems that can best be classified for Jupiter are light and dark; the light clouds usually are called zones, and the dark clouds bands.

On Jupiter we can see an equatorial zone on both sides of which there are the north and south tropical belts; these are followed by the north and south tropical zones, and next, closer to the poles, lie the north and south temperate belts; still more toward the poles the cloud structure is expressed less clearly and then disappears. In the north and south temperate zones it is possible to detect the presence of a poorly expressed belt; a southerly current also can be detected.

The problem of the nature of the cloud formations on Jupiter is associated closely with the problem of the chemical composition of its atmosphere. Hydrogen apparently is the principal substance. Methane and ammonia, which have been observed directly, are present in quantities of 150 and 7 atm-m, respectively. The most probable temperature of the cloud cover of Jupiter is -140°C (radiometric temperature), but in general the temperature ranges from -100 to -151°C (Ref. 25). These same values are obtained by radioastronomical measurements in the centimeter range. At such a temperature ammonia is in a solid state and methane is a gas. It has therefore been postulated that the cloud formations of Jupiter consist of an aerosol in which the most important role is played by crystalline ammonia (Ref. 25). It should be noted that the presence of gases such as methane and ammonia in the atmospheres of the planets of the Jupiter group affords great possibilities for all kinds of photochemical reactions. This is of more than a little importance in interpreting the relationship between various phenomena in the atmospheres of these planets and solar activity.

Returning to the problem of the cloud formations on Jupiter, we should note those which reveal a clear variability. This applies to the well-known red spot and the southern tropical disturbance. Both formations have a certain topographic affinity: both are situated in the south tropical zone. The following are the linear dimensions of the red spot: the length is about 50,000 km and the width is about 11,000 km (Ref. 25). The first information about it was published in the 18th and even the 17th century. It then was observed in the first half of the 19th century and has been observed regularly since 1878. It is now regarded as a long-lived atmospheric formation.

The change in the period of its rotation, characterizing the wind regime on Jupiter, is of the greatest interest. This will be discussed below. In general, the problem of the rotation of Jupiter is of very great interest. The equatorial and high-latitude regions of Jupiter rotate with different angular velocities; whereas the period of rotation of the equatorial zone is $9^h50^m30^s$, all of the other zones rotate with a different period, specifically, $9^h55^m40^s$. These of course are mean values, because, as we will see below, the periods of rotation of Jupiter do not remain constant, but change, and such changes reveal certain relationships to solar activity.

Such values, however, are used as a basis for defining systems for the reckoning of longitude. In system I the period of $9^h50^m30^s$ has been used, and in system II, $9^h55^m40^s$ (to be more exact, $40^s.632$ sec). Currently, as a result of the development of radioastronomical observations

of Jupiter, a need has arisen for a third system with a period of rotation equal to $9^h55^m29^s6$. Thus, the rotation of Jupiter reveals certain features in common with solar differential rotation, but at the same time it is impossible not to see the significant difference that the angular velocity of solar rotation decreases monotonically from the equatorial to the high latitudes, whereas on Jupiter the transition from more rapid to slower rotation occurs in a jump somewhere, near the Jovian latitude $\pm 15^{\circ}$.

It has not been possible to explain this distribution of the angular velocity of the rotation of Jupiter as a function of latitude, despite a number of attempts. In particular, we should mention here the investigations of Schoenberg (Ref. 26). He found it necessary to assume the presence on Jupiter of intrinsic heat sources, which in his opinion are of a volcanic nature. In general, however, the Jovian atmospheric circulation is an example of sharply expressed zonality, and this, to be sure, is entirely understandable if we take into account the great deflecting force (considerable rotational velocity). There is virtually no meridional motion on Jupiter.

In recent years Jupiter has been studied thoroughly by radioastronomical methods, which have made it possible to conclude that Jupiter has radiation belts (noticeable when Jupiter is investigated using relatively long waves) and apparently a strong magnetic field.

Saturn is very close to Jupiter in its physical properties. As on the largest planet of the solar system, Saturn has cloud belts, although the contrast is less clearly defined than on Jupiter. There are north and south tropical belts situated between $\pm 10^{\circ}$ and $\pm$ 20° latitude. The chemical composition of the atmosphere of Saturn is very close to the composition of the atmosphere of Jupiter, but the atmosphere has a somewhat lower temperature, according to Pettit, about $-160^{\circ}C$ (Ref. 27). This agrees well with radioastronomical data obtained at a wavelength of 3.4 cm. In this case the temperature is equal to $-167^{\circ}C$ and is capable of a more considerable freezing of ammonia, which is reflected in the photometric characteristics of the planet.

There are also transient cloud formations on Saturn, less stable than those on Jupiter, where the red spot has been observed for many decades with small interruptions, but a southern tropical disturbance, that covers almost a 110° longitude interval in the region of the south tropical zone, is retained virtually at all times and which only slightly changes in appearance. The bright or dark spots on Saturn are usually visible for one year and then disappear. The white spots that appeared in 1903 and 1933 are of the greatest interest. We note that in both cases this occurred at approximately the same point in the planet's orbit.

Also of interest is the dark spot situated in the north tropical zone in 1949-1950. It covered a rather large interval of longitude. The appearance of such spots, and also spectroscopic determinations, have made it possible to establish that there is a characteristic dependence of angular velocity on latitude for both Saturn and Jupiter. V. V. Sharonov has given the following summary (Ref. 25):

Latitude zone	Period of rotation
10°.	$10^{h}\ 12^{m}$
$10\text{-}20^{\circ}$	$10\ \ 15\text{-}10^{h}\ 20^{m}$
$35\text{-}40^{\circ}$	10 36-10 38
57°.	11 00-11 15

In all probability the pattern here is the same as on Jupiter, not on the sun. However, it still is premature to construct a model; further accumulation of data is necessary.

We will not discuss the characteristics of the atmospheres of Uranus and Neptune, first of all because information on these planets is even more limited than for Saturn, and secondly, as we will soon see, for the time being it is not possible to establish a manifestation of solar activity in the atmospheres of these planets for a variety of reasons.

Even in the middle of the last century the study of variations in the brightness of the planets had attracted the interest of investigators. It was assumed that by such a study it would be possible to detect variations of the solar constant (Ref. 28). Indeed this idea still has not been abandoned, and quite recently the problem was discussed in a paper by Johnson and Iriarte (Ref. 29). Observations of planetary brightness and its comparison with the phase of the 11-year solar cycle led to the following conclusion at the end of the last century (Ref. 30): the brightness values are slightly different in different years. Jupiter displays the clearest variations. They are observable somewhat less clearly on Mars, Saturn and Uranus.

The data serving as the basis for this investigation was supplemented and reworked in 1933 by W. Becker (Ref. 31). He discovered that Mars, Jupiter, Saturn, Uranus and Neptune change their brightness; this change in the cases of Mars and Saturn is irregular, for Jupiter it is with a period of 11.6 years, Uranus—8.4 years, and Neptune—21 years. In all these cases the amplitude of brightness variations was approximately the same—0.3 stellar magnitude. Thus, no data were

obtained on the relationship between planetary brightness and solar activity, except for Jupiter. No investigations of Venus were made. Observations of Venus certainly present certain specific difficulties. Nevertheless, since Venus is the planet with an atmosphere closest to the sun it is very probable that it is on Venus that it would be easiest to note the effect of active solar radiation.

Suspicion of the presence of Venusian auroras was expressed for the first time by Schoenberg (Ref. 32). In 1954 this problem was investigated by N. A. Kozyrev, who discovered the emission of molecular nitrogen λ 3914 and λ 4278 Å in the Venusian spectrum on its unilluminated side (Ref. 33). These bands also are observed in terrestrial auroral spectra. The problem of the molecular emission spectrum of Venus, which can be attributed to phenomena of the auroral type, later was studied by Newkirk (Ref. 34), who discovered other bands which also were noted by Kozyrev and which also can be related to luminescence of the upper layers of the atmosphere caused by active solar radiation (bands 4415 and 4435 Å).

Attempts to note manifestations of solar activity in the Jovian atmosphere, in addition to what has been said above concerning purely photometric investigations, were made at the beginning of this century. The fact of more turbulent changes in the atmosphere of Jupiter in some years than in others, established during the last century, requires us to postulate the presence of solar cycles. In actuality, cycles have been discovered which are suspiciously close to the 11-year cycle; Denning has obtained 9.8 years as the length of the cycle of alternation of years with particularly turbulent processes in the Jovian atmosphere; Williams has obtained 12 years and Wonaszek, 11.8 years (Refs. 35-37). We note that A. M. Bakharev (Ref. 38) made a comparison of variations in the intensity of the belts of Jupiter with intensifications of solar activity as indicated by Wolf numbers, and obtained a positive result. In this work he used the systematic comparison methods employed by O. V. Dobrovol'skiy in his study of the relationship between solar activity and brightness bursts of comets. It was taken into account that Wolf numbers for a terrestrial observer and for an imaginary observer on Jupiter will be different at the same physical time as a result of the difference in the heliocentric longitudes of the earth and Jupiter. Bakharev's positive results were obtained by taking into account corrections for this geometric effect. However, Bakharev's conclusion should be considered extremely preliminary.

An important investigation was made in 1948 by Hess and Panofsky (Ref. 39). These authors compared the periods of rotation of Jupiter on its axis, as indicated on the basis of the red spot, with the intensity of west-to-east transport of air masses in the earth's troposphere (the so-called index of zonal circulation) for 1900-1939. There was found to be a complete antiparallelism of the corresponding curves, obtained by the moving averages method. This result means that with an

intensification of west-to-east transport in the earth's atmosphere there is a shortening of the period of rotation of Jupiter, as established on the basis of the corresponding formations in its atmosphere. But a shortening of the period of rotation means an intensification of the wind in the direction of planetary rotation, that is, the result obtained is evidence that zonal circulation experiences parallel changes both on the earth and on Jupiter. It is obvious that solar activity is the only common factor that could account for such behavior of atmospheric circulation on two such widely separated planets.

The result obtained by Hess and Panofsky is evidence that when sufficient data are available much can be learned from a comparison of solar activity with phenomena in planetary atmospheres over long intervals of time—within the 11-year cycle and even longer. An additional argument in support of this point of view has been expressed by G. A. Shayn (Ref. 40) and Antoniadi (Ref. 41) in their discussions of the dependence of the rate of melting of the polar caps of Mars on the phase of the 11-year solar cycle. These authors have pointed out that the rate of melting increases parallel with an increase of sunspot number, increasing in years of maximum of the cycle and decreasing in the epoch of the minimum. The rate of melting of the polar caps is dependent on the intensity of the atmospheric circulation of this planet.

If we do not insist on the point of view that there are considerable variations of the solar constant in the 11-year cycle, and careful actinometric observations on the earth over a period of many years indicate that such variations do not occur, we must conclude that the rate of melting of the polar ice and snow of Mars should be dependent on the value of horizontal macroturbulent exchange of heat in the atmosphere of this planet. We already know that on earth this factor is related to solar activity; for example, it is sufficient to recall the investigations of L. A. Vitel's on the change of the number of well-developed cyclones with the phase of the 11-year cycle. There is an approximate parrallelism between variations of horizontal exchange and the change of vertical exchange in this cycle, which was investigated by A. A. Dmitriyev.

In view of the work of Hess (Ref. 42), who on the basis of radiometric observations and data on winds in the Martian atmosphere (obtained from observations of the movements of cloud formations in the atmosphere of that planet) drew a convincing picture of the distribution of pressure regions in its atmosphere, it can be stated that as on the earth cyclic variations of solar activity should lead to variations of heat exchange between the low and high latitudes.

We will now present the results of two investigations that we undertook to determine the manifestation of the long-term (i.e., 80-90 years) solar cycle in planetary atmospheres, especially in the atmosphere

of Jupiter. We have already had occasion to point out that no such investigations have been made in the past. One of the reasons for this is that the 80-90-year cycle has been reliably determined only quite recently, as we have seen in Chapter 1.

However, there is another factor which also was mentioned—the absence of reliable long uniform series of planetary observations. Nevertheless, some data can be selected here. These include, in particular, the observations of planetary brightness made by Becker. We have already mentioned this study, but now will discuss it in greater detail. Becker used observations made in the 19th and the beginning of the 20th centuries, including data published by G. Muller (Ref. 30), and reduced planetary brightness to a mean opposition and to a phase of 0^{o}.

The information was varied and was for the most part visual observational material of a type presenting considerable difficulties. It is well known what dangerous errors can be concealed in such information. Becker, a leading specialist in systems of stellar magnitudes, took careful steps to insure the constancy of the system to which planetary brightness was reduced. Since the brightness values for the entire epoch of the opposition were subjected to consideration, the data of each individual observer appears relatively more reliable. In the present century the observers themselves in most cases have given the brightness in the Harvard system of stellar magnitudes (abbreviated HRP). The observations made in the last century were reduced to the generally accepted system by substitution of the brightness value in the HRP system for the brightnes of the comparison star in place of the brightness used by the corresponding observer.

In the case of Müller's observations, it was necessary to use the color equation between the system used by this observer and the HRP system. Special methods were used to reduce King's photographic observations to the HRP system. We note that the reduction of Neptune's brightness involves certain difficulties associated with conversion from certain comparison stars at the beginning of the period of observations to other comparison stars at the end of this period. This circumstance makes data on the brightness of Neptune less reliable than data on the brightness of other planets. Most of the visual observations reduced by Becker required the use of photometers. Corrections for differential extinction, associated with a difference in the zenith distances of the star and planet, were introduced with the maximum possible accuracy for the corresponding epoch.

The gaps in the data cited by Becker are relatively small. It is only for the 1890's that there are no data on the brightness of the majority of the planets. In the remaining time the gaps rarely exceed 2-3 years in a row, and brightness data can be obtained by interpolation,

as was done graphically by Becker. Table 42 shows the years for which there are data concerning the brightness of various planets.

As already mentioned, the brightness data reduced to the HRP system were then reduced to the mean distance planet-sun and to the mean distance planet-earth, and also to the phase 0°. In the reduction of the brightness of Saturn, Becker also reduced the rings to such a position at which they disappear, that is, when they are turned edgewise to the observer.

In the tables of planetary brightness cited by Becker, data for one observer usually are given for each opposition. However, if the planet was observed by several individuals during a particular year, data are given for that observer who made the greatest number of observations with the minimum internal error. As a result, the mean errors and data on planetary brightness at one opposition have the values, as derived from Becker's reduced data, as shown in Tables 41 and 43. It is obvious that they must be regarded as somewhat too low, since such accuracy is possible only with an electrophotometer. The small values of the mean errors are associated with a formal increase of accuracy due to the large number of brightness determinations.

The principal conclusions drawn by Becker already have been cited. We also note that in addition to the already mentioned 8.4-year cycle in the brightness of Uranus, there is still another with a duration of 42 years, which is equal to half the time of the sidereal rotation of Uranus. This phenomenon is associated with the fact that the axis of Uranus as is well known lies in its orbital plane, and in the course of one revolution of Uranus around the sun will fall on the line of sight of a terresterial observer twice.

The disk of Uranus, possessing a rather considerable flattening, in this case will be projected on the celestial sphere in the form of an ellipse with a maximum area. It is obvious that planetary brightness will be maximal in such a position. It should be added that when the axis of Uranus coincides closely with the line of sight the observer sees the polar regions of this planet and virtually does not see its equatorial zone, crossed by dark belts. As a result of this position the brightness of Uranus in the corresponding epoch will be enhanced still further. A cycle with a period of 21.6 years has been discovered in the brightness variations of Neptune.

We formed moving 10-year means of brightness of Mars, Jupiter, Saturn and Uranus, and Wolf numbers (on the basis of mean annual values) in order to detect long-term variations in the brightness of the planets, and for their comparison with long-term variations of solar activity. Due to a disruption in the observations of the first three of these planets in the 1890's we separately analyzed the series from the middle

Table 41

Planet	Mean brightness error
Mars	±0.03
Jupiter	±0.07
Saturn	±0.03
Uranus	±0.03
Neptune	±0.04

Table 42

Mars	Jupiter	Saturn	Uranus	Neptune
1845	1846	1852	1864	1865
1846	1852	1858	1869	1879
1847	1857	1859	1874	1880
1849	1858	1860	1876	1882
1850	1862	1862	1877	1883
1852	1863	1863	1878	1884
1854	1865	1864	1879	1885
1856	1868	1865	1880	1886
1858	1869	1866	1881	1887
1860	1870	1868	1882	1895
1861	1871	1870	1883	1896
1865	1872	1872	1884	1897
1867	1873	1874	1885	1898
1869	1874	1876	1886	1899
1870	1876	1877	1887	1900
1871	1878	1878	1888	1901
1878	1879	1879	1894	1902
1880	1880	1880	1896	1908
1881	1881	1881	1898	1915
1882	1882	1882	1899	1920
1884	1883	1883	1900	1922
1886	1884	1884	1902	1923
1889	1886	1886	1908	1930
1901	1887	1887	1916	1931
1903	1890	1888	1917	1932
1906	1891	1891	1918	—
1908	1902	1906	1921	—
1909	1904	1907	1922	—
1914	1905	1908	1923	—
1916	1909	1909	1924	—
1918	1912	1910	1925	—
1920	1914	1916	1926	—
1921	1916	1917	1927	—
1922	1918	1918	1928	—
1927	1919	1920	1929	—
1929	1920	1921	1930	—
1931	1921	1922	1931	—
—	1922	1923	1932	—
—	—	1924	—	—

Table 43

Planet	Amplitude error	Mean brightness
Mars	0.40	±0.02
Jupiter	0.30	±0.03
Saturn	0.49	±0.02
Uranus	0.28	±0.02
Neptune	0.34	±0.02

of the 19th century to the 1890's and from the beginning of the current century to the last years for which observations are available. As noted in Chapter 1, if there is an 11-year cycle in observational data the formation of 10-year moving means almost entirely excludes this cycle. Thus, in the case of Jupiter, regardless of the cause of the 11-12-year cycle, in its brightness variations the forming of moving 10-year means greatly decreases the amplitude of this cycle.

We will now attempt to summarize certain results.

1. The amplitudes of long-term variations in planetary brightness as indicated by 10-year means have been cited in Table 43.

The amplitudes of the brightness variations of planets, therefore, are considerable and entirely real.

2. Figures 58 and 59, which show curves of the moving 10-year means of the brightness of Jupiter (curve 1, Figure 58) and Saturn (curve 1, Figure 59), as well as the moving 10-year means of the relative sunspot numbers (curves 2, Figures 58 and 59), show the following. The curve of long-term variations in the brightness of both Jupiter and Saturn, as indicated by moving 10-year means, is opposite to the curve of long-term variations of relative sunspot number, also based on moving 10-year means, determined for the same years for which data on planetary brightness were obtained.

The explanation for a certain difference in the curves of the moving 10-year means of relative sunspot number in Figures 58 and 59 is that the years for which data are available on planetary brightness do not coincide for the different planets. The form of the curve constructed on the basis of 10-year moving means is dependent on the initial year (i.e., the value) used as the beginning of the 10-year moving means.

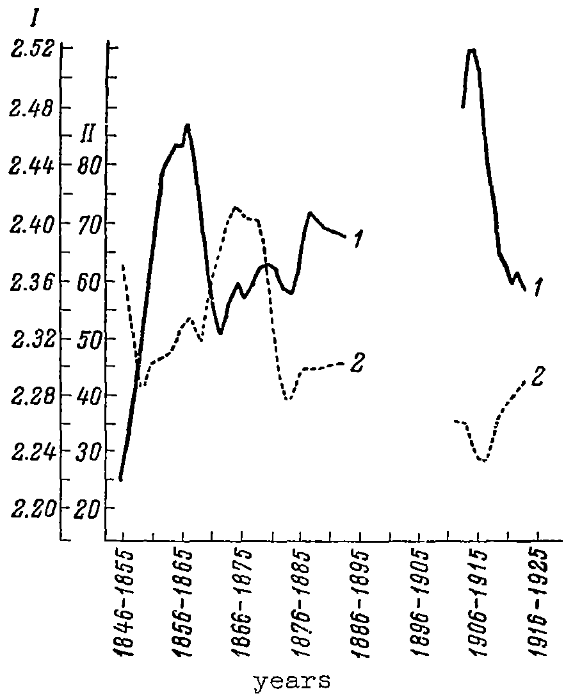


Figure 58. Brightness of Jupiter and solar activity. 1—10-year moving means of the brightness of Jupiter; 2—same, for Wolf numbers. Vertical scales: I—negative stellar magnitude; II—Wolf numbers

The curves of the relative sunspot number were constructed for the same years for which data are available on the brightness of the corresponding planets. This resulted in a certain difference in the shape of curves 2 in Figures 58 and 59.

One notes the great integral brightness of Jupiter and Saturn at the beginning of the current century in the epoch of the minimum of the 80-90-year cycle of solar activity, which, in particular, can be observed easily from curves 2 in Figures 58 and 59. On the other hand, both peaks of this cycle, occurring in the 1830's-1840's and in the 1860's-1870's, are accompanied by a low value of the integral brightness of Jupiter and Saturn.

3. It follows from Figure 60, which shows the curves of the 10-year moving means of the brightness of Mars (curve 1) and relative sunspot

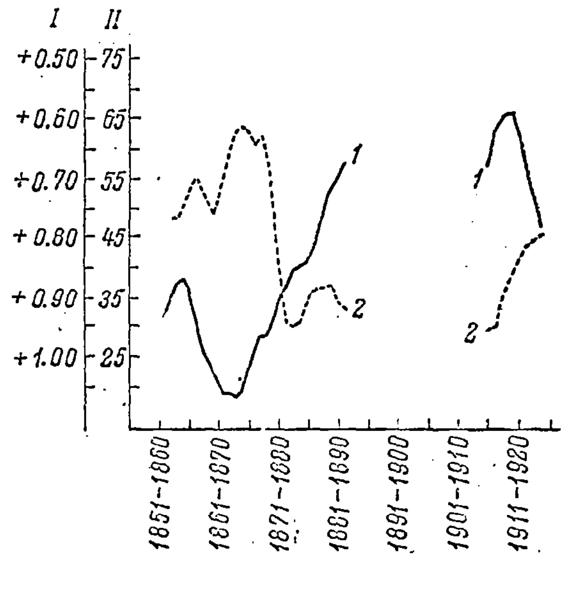


Figure 59. Brightness of Saturn and solar activity. 1—10-year moving means of the brightness of Saturn; 2—same, for Wolf numbers (the values I and II are the same as in Figure 58; the stellar magnitudes on the scale are positive)

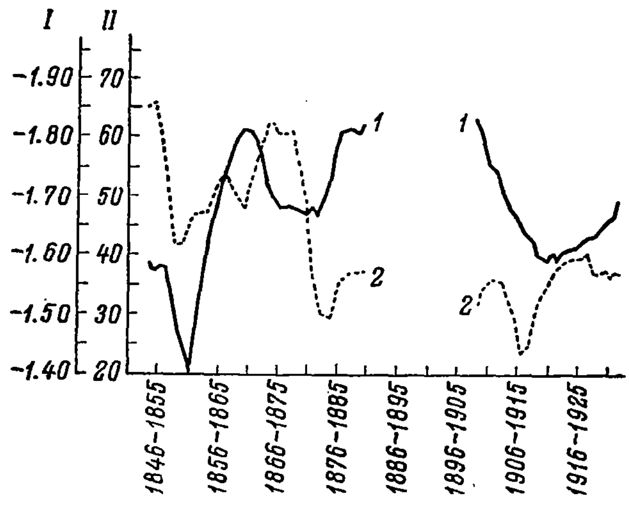


Figure 60. Brightness of Mars and solar activity. 1—10-year moving means of the brightness of Mars; 2—same, for Wolf numbers (the values I and II are the same as in Figure 58)

number (curve 2), that the long-term variations of the brightness of Mars are unrelated to the long-term variations of solar activity.

4. The relationship between the 10-year moving means of long-term variations of the brightness of Uranus (curve 1, Figure 61) and the similarly smoothed relative sunspot number (curve 2, Figure 61) cannot be established. Even if there is a relationship, it cannot be detected due to the presence of a 42-year period in the variations of brightness of this planet, which blurs out all other possible variations and as already mentioned, probably is caused by purely geometric factors. It is shown in Figure 61 that the formation of 10-year moving means of the brightness of Uranus clearly reveals a 42-year period in the variations of this brightness.

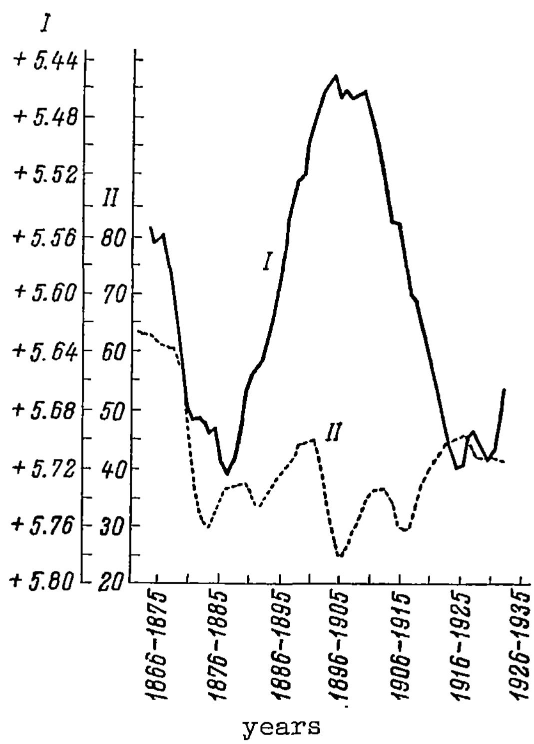


Figure 61. Brightness of Uranus and solar activity. 1—10-year moving means of the brightness of Uranus; 2—same, for Wolf numbers (the values I and II are the same as in Figure 58)

We will attempt, at least in the most general outlines, to give an interpretation of our results. We will first attempt to explain the opposite character of the variations in the brightness of Jupiter and Saturn and solar activity. However, we will first determine whether these results possibly are related to shortcomings in the photometric data used. Such an explanation hardly can be accepted. It cannot be assumed that the errors associated with certain small gaps in the series of observations would lead to such a close negative correlation between the long-term variations of the brightness of Jupiter and Saturn and solar activity. As we already have mentioned, the large gaps in the series of observations, such as the gap for Mars, Jupiter and Saturn in the 1890's, were excluded by the formation of separate series.

With respect to the possibility of long-term variations of extinction in the earth's atmosphere, even if this effect is real, it should be reflected to an identical degree on the brightness of both the planets and the comparison stars. The differential corrections caused by colorimetric factors undoubtedly are terms of the second order of magnitude and cannot lead to an amplitude of long-term variations of planetary brightness of several tenths of a stellar magnitude. This also applies to the possible effect of the different azimuths of Jupiter and Saturn, on the one hand, and bright comparison stars on the other. Moreover, it is not clear how this effect could lead to brightness variations opposite of the variations of relative sunspot number.

Attempts to attribute the results to long-term variations of the solar constant cannot be considered satisfactory. It is true that in recent years studies have appeared in which the authors attempt to demonstrate small variations of the solar constant during the 11-year cycle. We have in mind the already mentioned investigation of the brightness of Mercury and Neptune made by Johnson and Iriarte (Ref. 29). These photoelectric investigations revealed that during the period 1953-1958 the brightness of the mentioned planets changed to such an extent that it is possible to suspect an increase of the solar constant by 2 percent during these years. Abbot, Aldrich and Hoover (Ref. 43) discuss approximately the same amplitude of variation of the solar constant in the 11-year cycle. However, Sterne and Dieter found nothing of the sort when they analyzed these same data (Ref. 44). Johnson and Iriarte attribute the variation of the brightness of Uranus and Neptune, which they recorded, to intensifications of solar radiation in the violet part of the spectrum. However this may be, the long-term variations of the brightness of Jupiter and Saturn, revealing a relationship to the long-term variations of solar activity, are too considerable for them to be attributed to any admissible variation of the solar constant.

It must be concluded that the long-term variations of the brightness of Jupiter and Saturn are related to physical changes occurring in the atmospheres of these planets and dependent on long-term variations

of solar activity. Becker repeatedly points out in his study that numerous observers have indicated that changes in the brightness of planets are associated with changes on their surfaces. These facts are sound confirmation of the conclusion we have drawn that there is a relationship between variations in planetary brightness and physical changes on the planets themselves.

It is obvious that the absence of long-term variations of the planetary brightness of Mars as a function of solar activity should not lead us into confusion. The physical nature of Jupiter and Saturn is so different from that of Mars that it would be strange if these planets reacted in the same way in the 80-90-year cycle of solar activity. The physical changes which we observed on Jupiter and Saturn are changes in the atmospheres of these planets. The physical changes on Mars, however, are partially changes on its surface. With the earth as an example we know that solar activity is manifested far more clearly in the atmosphere than on the underlying surface. It is therefore completely understandable that it has not been possible to establish a relationship between long-term variations of the brightness of Mars and solar activity. As we already have seen, the uncertain results for Uranus can be attributed to unfavorable geometric conditions for the observations of this planet.

We will now consider the problem of the relationship between the variations of the brightness of Jupiter and solar activity within the 11-year cycle. The cycle of variations of the brightness of this planet, equal to 11.6 ± 0.4 year, established by Becker and N. P. Barabashev (Ref. 45) is close to the cycle of change of the colors of its equatorial zone; the cycle of the latter phenomenon is 11.95 years. The relative closeness of these cycles to the 11-year solar cycle and the period of the sidereal revolution of Jupiter around the sun (11.86 years) interested us in further investigation of the problem.

In order to be certain of the presence or absence of a relationship between variations of the brightness of Jupiter and the 11-year solar cycle it is necessary to make a quite rigorous evaluation of the observed deviations of the corresponding curves. One of the possible statistical approaches to the solution of such a problem is a clarification of what influence is exerted on the brightness of Jupiter by a particular phase of the 11-year cycle of solar activity. This problem can be clarified by the use of variance analysis.

This statistical method has the great advantage, in particular, that as it is based on the t distribution it can be applied to virtually any statistical sample, including one of small size. Variation analysis was used by O. V. Dobrovol'skiy (Ref. 4) to establish a relationship between the number of comets visible with the naked eye and solar activity in the 11-year cycle (see preceding section). Our computations were very close to those of Dobrovol'skiy.

The first step was to arrange the data on the brightness of Jupiter in accordance with the phase of the 11-year solar cycle. The phase of the cycle was computed using formula (6.7). The computation interval for the phases was 0.09, with the exception of the last class, in which the phase interval was 0.12 (from 0.88 to 0.99). This inequality of the classes is admissible and is taken into account by appropriate formulas. We have at our disposal almost 6.5 complete 11-year solar cycles (the total series of observations of the brightness of Jupiter includes seven 11-year cycles, but in the first and especially in the fifth of these cycles there are considerable gaps in the observations).

The following results are obtained when the data are arranged in accordance with the phase of the 11-year cycle and the formulas of variance analysis are used. The dispersion with respect to the phase of the cycle s_f^2 was 0.028 and the residual dispersion was $s_e^2 = 0.018$. Hence the value $T_H = 1.56$. Considering the number of degrees of freedom present in this case ($K_1 = 10$, $K_2 = 55$), in order to be able to speak of a real value of dispersion with respect to phase it is necessary that T_H, even in the case of a confidence coefficient of 0.01, would be 2.69. The θ test in this case gives 1, but as is well known, in order to be able to speak of the real value of the characteristic, the θ criterion should be > 3.

Finally, assuming that the distribution in the general sample is not normal, we obtain for the corresponding reality criterion a characteristic 0.82, which is less than 3. We can therefore speak with assurance of an absence of a relationship between the phase of the 11-year solar cycle and the values of the brightness of Jupiter at oppositions.

A different result is obtained in an investigation by the variation analysis method of the dependence of long-term variations of the brightness of Jupiter on long-term variations of solar activity. Arranging these same data in accordance with the number of the 11-year solar cycle, we find that the dispersion with respect to this criterion is $s_r^2 = 0.073$, and the residual dispersion is $s_e^2 = 0.014$. Therefore $T_H = 5.24$. In this case, in order to be able to speak of a dependence between the brightness of Jupiter and the number of the 11-year solar cycle with a confidence coefficient of 0.05, T_H should be not less than 2.25. In

the case of a confidence coefficient of 0.01 T_H should not be less than 3.13.

The use of the θ test also indicates the reality of long-term variations in the brightness of Jupiter—the criterion is equal to 6.52. Only by postulating an abnormal distribution in the general sample do we obtain 1.94 for the corresponding criterion. Thus, the reality of the relationship between the long-term variations of the brightness of Jupiter and solar activity is entirely probable.

We will now consider how the long-term variations of solar activity are manifested in the long-term variations of the periods of rotation of Jupiter on its axis, determined from the various cloud formations on this planet. In 1933-1934 A. St. Williams published several studies analyzing the variations of the periods of rotation of Jupiter on its axis, established from the spots of the various latitudinal currents.

The boundaries of these currents do not coincide fully with the boundaries of the corresponding cloud belts; the currents often extend over into the adjacent bright zone. For example, the spots lying in the north tropical zone and on the northern margin of the north equatorial belt form the north tropical current. Spots lying on the southern margin of the north equatorial belt and in the northern part of the equatorial zone form the north equatorial current.

Williams was able to demonstrate that periodic variations are detectable in the velocities of rotation of Venus, determined from the spots of certain currents. For the velocities of rotation established using the spots of the north tropical current the period is 12.4 years. The north and south equatorial currents are characterized by short cycles of velocities of about 2.5 years. There also are continuous changes of the velocity of rotation.

The observations of Williams and the data he used, published in the reports of the Jupiter Section of the British Astronomical Association, cover the period 1879-1928. During this time the velocities of rotation of the spots of the north equatorial current in general increased, but decreased in the south equatorial current. A decrease in velocity of rotation in these same years is also observed in the spots of the south temperate and south currents.

It should be noted that Williams was unable to explain the periods of 2.5 and 12.4 years that he found, nor what he called the secular change of the velocity of rotation of spots on Jupiter, that is, the continuous change of the velocities of the spots of both equatorial currents, the south temperate current, and the south current on this planet.

Williams knew only of the 11-year solar cycle and the period of sidereal revolution of Jupiter of 11.86 years as periods or cycles having a physical basis.

We have used data on the periods of rotation of Jupiter on its axis determined from the spots of different latitudinal currents for 1900-1939. We could not use the earlier data of Williams because of considerable gaps in his observations in the 1890's. We formed 10-year moving means of the periods of rotation of five different Jovian currents at each of its oppositions. In addition to the north tropical, north and south equatorial, south temperate and south currents we also considered the long-term variations of the rate of rotation of Jupiter, determined from the red spot and the south tropical disturbance. These velocities were analyzed by the same method as the velocities of the latitudinal currents.

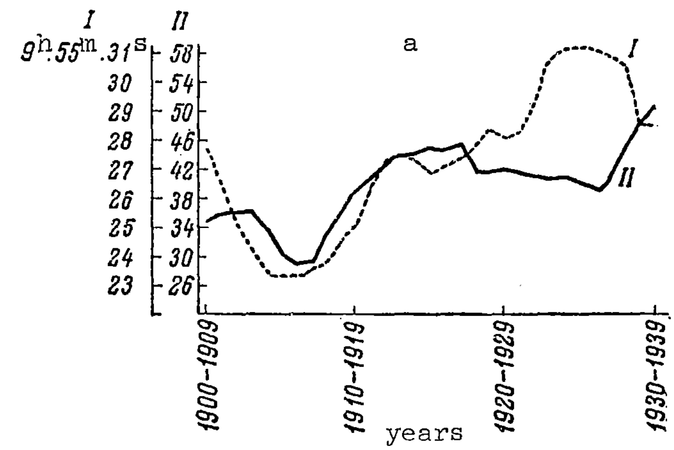
Figure 62 shows the curves of the 10-year moving means of the periods of rotation of Jupiter.

The curves applying to Jupiter have been denoted I. The figure II is used to denote the curve of the 10-year moving means of relative sunspot number from the decade 1900-1909 through to the decade 1930-1939. Table 44 gives mean data on the periods of rotation, their amplitudes during the epoch considered, errors in determination of periods of rotation and their amplitudes, etc. A study of Figure 62 and Table 44 makes possible the following conclusions.

1. Long-term variations of the periods of rotation of Jupiter on its axis, determined from spots in different latitudinal currents, and apparently also from the red spot and the south tropical disturbance, are completely real. The values of the amplitudes of the variations of these periods can be considered reliable.

2. Within the limits of the considered time interval (1900-1939), coinciding with the ascending branch of the 80-90-year cycle of solar activity, it is possible to establish two types of long-term variations of the periods of rotation of Jupiter:

(a) long-term variations that in general correlate positively with long-term variations of solar activity; these include periods of rotation established from the spots of the north tropical, south equatorial, south temperate and south currents, and the south tropical disturbance; in all of these cases the higher the long-term level of solar activity, the longer is the period of rotation of Jupiter, as determined from the corresponding spots, and therefore the smaller is the linear velocity of rotation of this planet at the corresponding latitude and level;



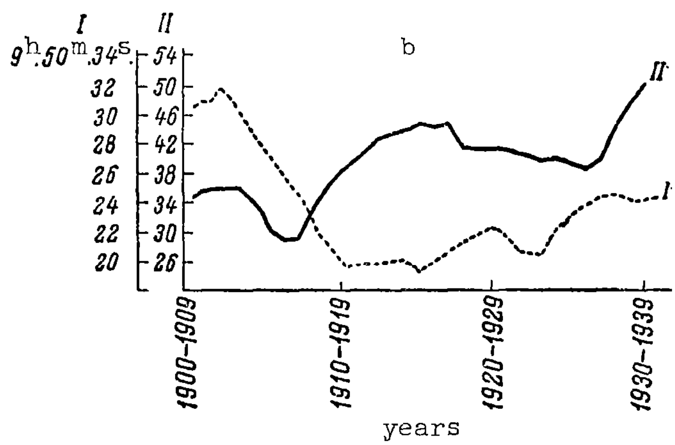


Figure 62. Dependence of the rotation of Jupiter on solar activity. I—periods of rotation of Jupiter; II—Wolf numbers; a—for north tropical current; b—for north equatorial current; c—for south equatorial current; d—for red spot; e—for south tropical disturbance; f—for south temperate current; g—for south current (continued)

(b) long-term variations which correlate negatively with long-term variations of solar activity; these include the north equatorial current and to a certain degree the red spot. In these cases the higher the level of solar activity, the shorter is the period of rotation of Jupiter, determined from the corresponding spots, and therefore the greater is the linear velocity of rotation of this planet at the corresponding latitude.

Table 44

Current or object	Latitudinal limits (1900-1928)	Mean period of rotation (1900-1939)	Mean error, sec.	Maximum period					
				Initial values			10-year moving mean		
				year	period, sec.	mean error, sec.	decade	period, sec.	mean error, sec.
North tropical current	from +10 to +20°	$9^h55^m27.4^s$	±1.9	1932	36.0	±1.0	1925-1934	31.2	<±1
North equatorial current	from + 0 to +10°	$9^h50^m24.6^s$	±2.5	1902	36.0	>±5	1902-1911	32.0	<±1
South equatorial current	from 0 to -16°	$9^h50^m37.4^s$	±4.0	1932	63.6	±3.5	1915-1924	49.6	±1
Red spot	-20°	$9^h55^m38.5^s$	—	1939	44.5	—	1900-1909	40.9	—
South tropical disturbance	-23°	$9^h55^m29.2^s$	—	1935	39.0	—	1930-1939	36.1	—
South temperate current	from 27 to -37°	$9^h55^m19.7^s$	±3.3	1937	25.0	>±5	1925-1934	22.2	<±1
South current	from -37 to 55°	$9^h55^m6.8^s$	±5	1938	11.0	>±5	1924-1933	9.3	<±1

Table 44 (continued).

	Minimum period					Amplitude							
	initial values		10-year moving mean			initial values				10-year moving mean			
year	period, sec.	mean error, sec.	decade	period, sec.	mean error, sec.	seconds	mean error, sec.	km/hour	mean error, sec.	seconds	mean error, sec.	km/hour	mean error, sec.
1938	13.0	±1.0	1906-1915	23.4	±1	23.0	±1.4	27.7	±1.7	27.8	±1	9.4	+1.2
1913	12.1	±1.0	1915-1924	19.8	±1	23.9	±5.1	28.7	±6.1	12.2	±1	14.6	±1.2
1913	11.6	±5	1907-1916	25.4	1	52.0	±6.1	62.5	±7.3	24.2	±1	29.1	±1.2
1924	32.4	—	1916-1925	35.6	—	12.1	—	14.5	—	5.3	—	6.4	—
1902	16.0	—	1901-1910	20.6	—	23.0	—	27.6	—	15.5	—	18.6	—
1911	12.7	±3.5	1915-1924	17.9	$<\pm1$	12.3	±6.1	12.7	$>\pm6.3$	4.3	$<\pm1$	4.4	$<\pm1.03$
1908	1.0	±3.5	1904-1913	4.0	$<\pm1$	10.0	$>\pm6.1$	10.3	$>\pm6.3$	5.3	$<\pm1$	5.5	±1.03

Note: The latitudes are approximate for the red spot and the south tropical disturbance; there are no data on the errors of determination of the periods of rotation.

curves diverge noticeably. Beginning with the decade 1915-1924 to the last of the decades considered, 1930-1939, both curves are antiparallel to one another. This resembles the variation of certain indices of circulation of the earth's atmosphere. There is a certain analogy between the curve of periods of rotation of Jupiter, determined from the spots of the north tropical belt and the curve of these periods, determined from the spots of the south belt (Figure 62g).

5. It is characteristic that on the basis of the five latitudinal currents and formations in the Jovian southern hemisphere the periods of rotation of the planet, determined from the spots of the four currents, correlated positively with the long-term variations of solar activity. Only the periods of rotation determined from the red spot correlated negatively with the long-term variations of solar activity.

6. The 10-year moving means of the periods of rotation of Jupiter, determined from the spots of the south temperate current, in general, as already mentioned, correlated positively with the long-term variations of solar activity. However, with respect to details one immediately notes the opposite character of the two curves (Figure 62f) during the considered interval of time.

7. Variations of the periods of rotation determined from the spots of the south equatorial current (52 sec on the basis of initial values and about 25 sec on the basis of 10-year moving means) reveal the maximum amplitude during the considered time. Second place is taken by the south tropical disturbance (23 sec according to initial data and about 15 sec according to 10-year moving means). Third place is held by the north equatorial current (23.9 sec according to initial values and about 12 sec according to 10-year moving means). Fourth place is occupied by the north tropical current (23 sec according to initial data; about 8 sec according to 10-year moving means). Fifth place is occupied by the red spot and the south temperate current (12.1 and 12.3 sec according to initial values, and 5.5 and 4.5 sec according to 10-year moving means, respectively). Last place is occupied by the south current (10.0 sec according to initial values and about 5 sec according to 10-year moving means).

If we average the amplitudes of the north and south equatorial currents, and also the north tropical current and the south tropical disturbance, the first place with respect to the amplitude of long-term variations of the periods of rotation will be occupied by the spots of the equatorial currents. Second place is occupied by the spots of tropical formations. The amplitude of the variations of the periods of rotation of Jupiter as determined from the red spot has the same order of magnitude as the amplitude as determined from the high-latitude currents of the souther hemisphere, although the red spot is situated at the same latitude as the south tropical disturbance (about -20°).

8. One notes a coincidence of the year when the period of rotation of Jupiter as determined from the spots of the north equatorial current was minimal with the year of the minimum period of rotation, as determined from the spots of the south equatorial current (1913). In exactly the same way there is a coincidence of the years with maximum periods of rotation of Jupiter as determined from the spots of the north tropical and south and equatorial currents (1932).

In completing the consideration of the problem of long-term variations of the periods of rotation, determined from different formations in its atmosphere, we should mention the results obtained in the averaging of the periods of rotation of this planet, determined from the red spot, over a long series of years. As already noted, the red spot was discovered for the first time in the second half of the 17th century. Then the red spot was not observed during a large part of the 18th century and the beginning of the 19th century. Observations did not begin until 1831. Table 45 gives the periods of rotation of Jupiter, determined from the red spot, averaged over a long series of years; and the relative sunspot number, averaged for 50-year periods. Table 45 shows that with an increase of the long-term mean relative sunspot number the period of rotation of Jupiter, determined on the basis of its red spot, decreases. Table 45 indicates the character of the manifestation of solar cycles of longer duration in the atmosphere of Jupiter.

In summary, it must be noted that at the present time it would be premature to interpret the derived facts. With respect to planetary brightness, we therefore limit ourselves to those considerations discussed above. With respect to the periods of rotation of Jupiter, or to the point, with respect to the winds on Jupiter, for the time being it unfortunately is impossible to say more than has been said with respect to planetary brightness, in fact, we can only say less.

Present-day concepts concerning the Jovian atmospheric circulation are extremely modest. We still do not even know the source of energy of this circulation. Under these conditions any attempt to interpret variations in such circulation, that is, to a certain degree of an effect of a higher order, is premature. This interpretation is a matter for the future, and will be possible the sooner thorough, systematic observations of Jupiter and the other planets are organized.

We would like to emphasize that the study of the exceedingly interesting and very promising problem of the manifestation of solar activity on the planets is only beginning. For now there are no specific instruments for use in solving the problems outlined here, and research methods have not been properly developed. It must be assumed that in the immediate future the situation will change, and a new field of astrophysics, important in both practical and theoretical respects, will be developed.

Table 45

Fifty-year period	Period of rotation of Jupiter, determined from the red spot (averaged data)	Relative sun-spot number (averaged)	Remarks
1664-1713	$9^h 55^m 55.1^s$	14.1	Period of rotation determined from red spot, averaged for the years 1664-1672 and 1713. Relative sunspot number from 1700 to 1713. No data on mean annual values of relative spot numbers before 1700. However, it is known that at the end of the 17th century solar activity in any case was not higher than at the beginning of the 18th century. Thus, the number 14.1 is in fact the upper limit.
1714-1763	no data	40.2	
1764-1813	no data	47.2	
1814-1863	$9^h 55^m 35.4^s$	48.1	Period of rotation for years 1831-1863.
1864-1913	$9^h 55^m 37.4^s$		Period of rotation for years 1864-1889.

C

APPENDIX

Tables of Certain Solar and Geophysical Indices

Explanation of Tables

Table I gives data on the indices $\bar{a}$ and $\bar{b}$, introduced by M. S. Eygenson (Refs. 1, 2). The index $\bar{a}$ is the mean lifetime of all spot groups, expressed in solar rotations, and is derived by division of all groups (the second column in Table I, if groups with a lifetime of only one solar rotation are taken into account, and the third column if such short-lived groups are not taken into account) by the physical number of groups. In other words, the number of groups was weighted for the number of their repeated appearances. The index is derived for a year. The value $\bar{a}$ represents the value a smoothed, using a three-term formula. The index $\bar{b}$ is the mean duration (lifetime) of repeating spot groups, expressed in solar rotations. This value is derived by division of the number of appearances of repeating groups by the number of such groups.

Table II gives the index: mean group area for a year or AGA. The index was computed using data from the Greenwich catalog separately for all groups with a mean area (for an appearance) of $>$ 100 millionths of a solar hemisphere, and for the same groups with a mean area of $>$ 300 millionths of a solar hemisphere.

Table III gives the index S_M or MGA (mean annual value of the maximum areas for appearance of a group). Data on areas of the groups also were taken from the Greenwich catalog. The index was proposed by T. L. Mandrykina (Ref. 3).

Table IV contains the Kopecky indices F_0 and T_0. The index F_0 gives the number of groups forming on both the visible and far sides of the solar disk. Kopecky developed special methods for determining how many groups formed during a particular period of time on the far side of the disk. The index T_0 represents the mean lifetime of a group in days for a particular period of time. Each index was used in forming moving averages for 27 solar rotations (for approximately two years). The mean values then were computed for each year (Ref. 4).

Table V gives the Bartels index δW_1, determined in the following manner:

$$\delta_1 W_1 = \omega_1 (A_s - A_{s,\,R=50}).$$

Here A_s is the excess of the mean value of the horizontal component of the geomagnetic field at the Huancayo Observatory (in Peru, at the magnetic equator) at 75°W from 0900 to 1400 hours over its value at the same observatory between 0000 and 0500 hours; $A_{s,\ R\ =\ 50}$ is the same value when the Wolf number is 50. This is used as the normal value. In determination of the mean monthly values A_s was used in the computations only for days when the international magnetic characteristic C was less than 1.2. The factor ω_1 has an annual variation and is taken from a table compiled with the argument $A_{s,\ R\ =\ 50}$, which is different for different months of the year (Ref. 5). The value δR represents the deviation of Wolf numbers in a system compared with the δW_1 system. The value δR is derived using the simple formula:

$$\delta R = 0.649\ (R-50),$$

where R are smoothed mean monthly Wolf numbers.

Table VI gives the mean annual values of the Bartels index P, characterizing solar corpuscular radiation. The index is computed using the formula

$$\delta P = f U_1 - g,$$

where f and g are tabulated for each month of the year. The values u_1 represent the values of the U measure, converted to another scale:

$$U = \frac{U\ (H)}{\sin\theta \cos\psi},$$

where U(H) is the absolute difference in the values of the horizontal component of the geomagnetic field on the particular and preceding days (Bartels used the data for the Potsdam Magnetic Observatory), θ and ψ are geomagnetic latitude and declination, respectively. The conversion of u to u_1 is possible by use of a special table.

Table I

Year	Index $\bar{a}$, with groups having a lifetime of only one rotation	Index $\bar{a}$, without groups having a lifetime of only one rotation	Index $\bar{b}$	Year	Index $\bar{a}$, with groups having a lifetime of only one rotation	Index $\bar{a}$, without groups having a lifetime of only one rotation	Index $\bar{b}$
1875	1.160	1.216	2.143	1915	.069	.132	.280
1876	.045	.060	.500	1916	.077	.125	.194
1877	.167	.220	.333	1917	.133	.179	.288
1878	.000	.000	—	1918	.138	.216	.208
1879	.100	.154	.000	1919	.140	.219	.528
1880	.160	.205	.167	1920	.096	.169	.235
1881	.158	.228	.100	1921	.136	.240	.316
1882	.159	.208	.273	1922	.139	.246	.429
1883	.160	.209	.600	1923	.016	.031	.000
1884	.184	.230	.464	1924	1.070	1.119	2.333
1885	.128	.159	.333	1925	.110	.169	.538
1886	.167	.204	.800	1926	.111	.177	.156
1887	.113	.113	.333	1927	.092	.124	.417
1888	.080	.100	.500	1928	.138	.198	.333
1889	.231	.273	.200	1929	.101	.155	.416
1890	.018	.026	.000	1930	.082	.123	.375
1891	.129	.166	.280	1931	.064	.096	.000
1892	.127	.186	.364	1932	.087	.146	.375
1893	.131	.170	.314	1933	.163	.304	.500
1894	.126	.181	.211	1934	.082	.122	.000
1895	.130	.205	.324	1935	.117	.168	.273
1896	.096	.140	.063	1936	.108	.156	.121
1897	1.104	1.154	2.546	1937	.068	.088	.160
1898	.094	.128	.112	1938	.106	.136	.226
1899	.056	.075	.333	1939	.121	.157	.147
1900	.119	.128	.000	1940	.068	.083	.000
1901	.071	.100	.000	1941	.071	.084	.364
1902	.028	.056	.000	1942	.127	.154	.200
1903	.095	.130	.091	1943	.191	.231	.500
1904	.088	.124	.048	1944	.070	.083	.250
1905	.160	.235	.294	1945	.059	.071	.000
1906	.147	.211	.387	1946	.091	.106	.107
1907	.209	.268	.379	1947	.119	.133	.325
1908	.150	.192	.480	1948	.146	.167	.395
1909	.167	.212	.350	1949	.140	.173	.488
1910	.121	.172	.222	1950	.164	.273	.724
1911	.033	.046	.000	1951	.130	.165	.579
1912	.054	.083	.000	1952	.108	.333	.333
1913	.000	.000	—	1953	.216	.333	.333
1914	.045	.076	.000				

Table II

Year	AGA (all groups)	AGA (>100)	AGA (>300)	Year	AGA (all groups)	AGA (>100)	AGA (>300)
1879	89	154	0	1907	136	271	490
1880	165	260	477	1908	108	244	493
1881	133	219	387	1909	125	245	385
				1910	86	226	435
1882	148	317	565	1911	58	139	0
1883	171	278	559	1912	72	152	0
1884	136	250	450	1913	40	0	0
1885	133	240	380	1914	91	296	448
1886	126	262	476	1915	90	259	519
1887	96	258	382	1916	82	236	494
1888	81	247	394	1917	112	252	488
1889	106	259	318	1918	97	231	403
1890	91	226	440	1919	119	256	487
1891	121	232	466	1920	112	244	649
1892	135	273	491	1921	114	253	429
1893	118	241	454	1922	120	321	544
1894	119	272	511	1923	61	0	0
1895	119	226	403	1924	124	313	498
1896	124	293	495	1925	121	259	488
1897	134	273	549	1926	140	351	631
1898	113	273	450	1927	108	264	523
1899	71	245	392	1928	131	286	553
1900	83	168	0	1929	132	270	459
1901	88	211	311	1930	96	264	496
1902	117	249	394	1931	90	233	435
1903	103	217	448	1932	79	204	379
1904	95	239	399	1933	107	—	471
1905	153	282	558	1934	79	203	606
1906	105	249	481	1935	133	268	516

Table III

Year	MGA	Year	MGA	Year	MGA	Year	MGA
1900	133.6	1916	91.2	1932	91.3	1948	196.8
1901	102.3	1917	135.9	1933	102.9	1949	200.3
1902	118.2	1918	116.9	1934	98.3	1950	191.1
1903	125.3	1919	130.1	1935	158.4	1951	187.1
1904	123.2	1920	118.7	1936	137.2	1952	140.2
1905	181.5	1921	120.8	1937	179.1	1953	103.1
1906	136.8	1922	131.0	1938	165.6	1954	64.0
1907	180.9	1923	59.9	1939	174.4	1955	160.6
1908	140.1	1924	126.4	1940	176.5	1956	228.0
1909	175.9	1925	140.7	1941	153.8		
1910	110.4	1926	149.4	1942	145.4		
1911	70.7	1927	148.0	1943	190.2		
1912	87.8	1928	161.6	1944	95.9		
1913	67.2	1929	148.3	1945	138.7		
1914	83.4	1930	115.6	1946	223.2		
1915	88.5	1931	119.4	1947	226.6		

Table IV

Year	F_0	T_0	Year	F_0	T_0
1875	93.28	10.77	1913	11.36	14.57
1876	74.74	8.81	1914	185.74	6.20
1877	55.68	11.04	1915	948.05	5.03
1878	18.00	10.38	1916	1057.05	5.84
1879	37.57	7.68	1917	917.70	9.58
1880	174.24	10.77	1918	978.93	7.83
1881	415.67	8.29	1919	736.16	8.14
1882	386.64	10.51	1920	539.66	6.62
1883	351.68	12.61	1921	375.00	6.26
1884	417.28	11.92	1922	265.63	5.74
1885	352.23	10.64	1923	113.96	5.89
1886	168.96	10.77	1924	175.10	8.60
1887	124.80	8.44	1925	539.70	8.36
1888	62.22	10.16	1926	770.66	7.40
1889	36.08	11.80	1927	694.22	8.60
1890	69.54	10.16	1928	765.81	8.88
1891	281.60	10.77	1929	768.26	8.14
1892	614.64	9.13	1930	393.12	9.67
1893	761.76	10.06	1931	207.64	10.51
1894	824.86	8.52	1932	184.86	6.99
1895	691.06	7.90	1933	105.64	5.23
1896	373.80	8.36	1934	145.04	5.89
1897	307.19	7.68	1935	311.10	10.16
1898	238.96	8.60	1936	743.05	9.40
1899	114.40	8.44	1937	885.76	11.14
1900	52.17	14.76	1938	996.30	8.67
1901	27.36	7.33	1939	698.88	9.49
1902	56.98	5.89	1940	481.25	10.90
1903	222.48	8.60	1941	303.18	11.92
1904	416.12	8.60	1942	252.83	9.40
1905	529.93	9.13	1943	136.32	8.22
1906	479.15	9.96	1944	108.29	7.68
1907	508.28	9.31	1945	238.74	11.14
1908	482.56	8.44	1946	537.84	12.04
1909	345.71	10.27	1947	808.08	12.76
1910	245.70	8.36	1948	705.33	13.22
1911	101.68	6.38	1949	747.34	12.46
1912	78.68	5.13	1950	429.87	12.17

Table V

Year	January δR	January δW₁	February δR	February δW₁	March δR	March δW₁	April δR	April δW₁
1922	—	—	—	—	3	−11	−25	−31
1923	−30	−28	−31	−22	−30	−22	−29	−11
1924	−32	−27	−29	−27	−31	−28	−25	−24
1925	−29	−13	−18	−29	−21	−20	−12	−11
1926	14	10	13	16	8	3	− 8	9
1927	21	10	28	22	13	23	29	27
1928	22	14	16	8	23	15	20	28
1929	12	13	9	24	0	11	2	13
1930	10	22	− 1	8	−10	8	− 8	6
1931	−23	−20	− 5	−19	−13	−18	−12	−17
1932	−25	−33	−25	−28	−25	−17	−25	−36
1933	10	22	− 1	8	−18	8	− 8	6
1934	−31	−32	−27	−33	−30	−26	−25	−26
1935	−20	−25	−19	−15	−18	−30	−25	−27
1936	8	13	16	16	18	8	16	7
1937	53	53	51	56	22	32	38	24
1938	31	22	45	21	23	17	33	37
1939	19	30	18	24	+10	0	+38	+24

Table V (continued)

Year	May δR	May δW₁	June δR	June δW₁	July δR	July δW₁	August δR	August δW₁
1922	−27	−23	−29	−19	−25	−38	−29	−32
1923	−31	−33	−27	−27	−30	−20	−32	−18
1924	−19	−20	−17	−18	−14	−29	−20	−21
1925	− 5	−16	− 1	0	− 8	− 9	− 8	−17
1926	9	23	16	7	1	− 2	8	14
1927	19	13	6	12	3	2	3	4
1928	18	34	27	38	31	45	22	31
1929	5	7	14	4	13	18	10	5
1930	− 8	0	−14	0	−18	− 4	−16	−13
1931	−16	−29	−23	−31	−21	− 5	−24	−21
1932	−21	−20	−18	−19	−26	−25	−28	−38
1933	−30	−33	−29	−33	−30	−47	−32	−26
1934	−19	−20	−28	−19	−27	−29	−27	−30
1935	−15	−13	− 3	−16	−10	− 7	−13	− 4
1936	3	21	13	20	1	13	24	1
1937	43	40	52	48	62	52	57	60
1938	50	31	31	19	75	48	45	30
1939	44	37	33	38	31	50	36	47

Table V (continued)

Year	September		October		November		December		Mean annual	
	δR	δW_1	δR	δW_1	δR	δW_1	δR	δW_1	δR	δW_1
1922	—29	—27	—29	—27	—28	—26	—21	—28	—29	—30
1923	—24	—23	—25	—18	—26	—28	—30	—34	—31	—25
1924	—16	—15	—16	—14	—18	—14	—22	— 9	—23	—22
1925	6	10	12	8	6	8	32	20	— 4	— 6
1926	7	7	14	— 4	6	10	19	5	9	9
1927	12	11	8	20	11	3	— 3	9	14	9
1928	26	29	7	27	0	4	6	22	19	26
1929	—10	3	3	4	20	9	38	29	10	12
1930	—12	—18	—10	—20	— 9	— 7	—16	—25	—10	— 4
1931	—20	—26	—26	—21	—20	—24	—21	—25	—20	—53
1932	—30	—29	—27	—34	—27	—41	—25	—35	—27	—32
1933	—29	—28	—30	—29	—32	—32	—32	—25	—31	—33
1934	—30	—28	—29	—26	—27	—19	—23	—18	—29	—27
1935	— 5	—15	2	—13	9	1	8	13	—10	—13
1936	17	24	25	31	42	36	47	33	20	20
1937	33	35	49	32	16	40	25	33	45	45
1938	26	24	32	32	47	35	28	42	41	32
1939	41	27	25	30	—	—	—	—	32	33

Table VI

Year	δP	Year	δP	Year	δP	Year	δP	Year	δP
1872	56	1886	— 3	1900	—31	1914	—33	1928	13
1873	15	1887	—15	1901	—46	1915	3	1929	21
1874	11	1888	—17	1902	—45	1916	5	1930	17
1875	—15	1889	—24	1903	—11	1917	28	1931	—14
1876	—31	1890	—35	1904	—15	1918	32	1932	—17
1877	—21	1891	— 1	1905	1	1919	39	1933	—21
1878	—32	1892	46	1906	—10	1920	26	1934	—21
1879	—31	1893	20	1907	7	1921	8	1935	— 7
1880	— 8	1894	39	1908	10	1922	—14	1936	12
1881	4	1895	10	1909	13	1923	—22	1937	47
1882	27	1896	5	1910	— 5	1924	—12	1938	56
1883	8	1897	— 4	1911	—22	1925	0	1939	44
1884	12	1898	— 7	1912	—36	1926	33	1940	41
1885	8	1899	—28	1913	—48	1927	17	—	

REFERENCES

Chapter 1

1. Horrebow, P. Elementa Mathesos I, Copenhagen, 1732.
2. Gleissberg, W. Häufigkeit der Sonnenflecken (Frequency of Sunspots). Akademie Verlag, Berlin, 1952.
3. Waldmeier, M. Astronomische Mitteilungen, No. 221, 1959.
4. Kuiper, G. Solntse (The Sun). Moscow, IL, 1957.
5. Waldmeier, M. Ergebnisse und Probleme der Sonnenforschung (Results and Problems of Solar Investigations). Leipzig, Geest und Portig, 1955.
6. Kopecky, M. BAC, Vol. 9, 35, 1958.
7. Kleczek, J. Publications of the Central Astronomical Institute, Prague, No. 22, 1952.
8. --- Publication of the Astrophysical Observatory, Ondrejov, No. 24, 1953.
9. Minnis, C. M., Bazzard, G. H. Journal of Atmospheric and Terrestrial Physics, Vol. 14, 214, 1959.
10. Appleton, E. V. Philosophical Magazine, (7), 24, 144, 1939.
11. Waldmeier, M. Astr. Mitt., No. 133, 1935.
12. --- Experientia, Vol. 1, 354, 1946.
13. Ludendorf, H. Zeitschrift für Astrophysik, Vol. 3, 370, 1931.
14. Xanthakis, J. Annals of Astrophysics, Vol. 22, 85, 1959.
15. Chistyakov, V. F. Byulleten' VAGO, No. 15, 1954.
16. Ol', A. I. Byulleten' KISO, No. 1, 1949.
17. Vitinskiy, Yu. I. Prognozy solnechnoy aktivnosti (Forecasts of Solar Activity). Moscow-Leningrad, Izd-vo AN SSSR, 1963.
18. Kozik, S. M. Astronomicheskiy Zhurnal, Vol. 26, 28, 1949.
19. Kopecky, M. BAC, Vol. 2, 176, 1951.
20. Wolf, R. Handbuch der Astronomie, Zürich, Schulthiss, Vol. 2, 1892.
21. Kimura, H. MN, Vol. 73, 543, 1913.
22. Oppenheim, S. AN, Vol. 232, 369, 1928.
23. Stewart, Z. Q., Eggleston, F. C., Panofsky, H. A. A. Ap. J., Vol. 88, 385, 1938.
24. Gleissberg, W. Ap. J., Vol. 110, 94, 1949.
25. --- Z. Astrophys., Vol. 18, 199, 1939.
26. Chvojkova, E. BAC, Vol. 3, 21, 1952.
27. Wolf, R. Wien Kortweg Sitzungsberichte, 1883.
28. Turner, H. MN, Vol. 74, 82, 1913; Vol. 85, 467, 1925.

29. Hale, G. E., Nicholson, S. Ap. J., Vol. 62, 270, 1925.
30. Gnevyshev, M. N., Ol', A. I. Astron. Zh., Vol. 25, 18, 1948.
31. Chistyakov, V. F. Byull. VAGO, No. 25, 1959.
32. Wolf, R. Astr. Mitt., No. 14, 1862.
33. Waldmeier, M. Z. Astrophys., Vol. 43, 149, 1957.
34. Wolfer, A. Astr. Mitt., No. 100, 1909.
35. Eygenson, M. S. Astron. Zh., Vol. 25, 101, 1948.
36. --- Tsirkulyar LAO, No. 34, 1958.
37. Stetson, H. T. Sunspots in Action. Ronald, N. Y., 1947.
38. Dizer, M. Ann. astrophys., Vol. 19, 207, 1956.
39. Balli, E. Ann. astrophys., Vol. 18, 118, 1955.
40. Waldmeier, M. Z. Astrophys., Vol. 43, 149, 1957.
41. Gleissberg, W. Publications of Istanbul University Observatory, No. 26, 1944.
42. --- Observatory, Vol. 62, 158, 1939.
43. --- Observatory, Vol. 63, 215, 1940.
44. --- Ap. J., Vol. 96, 234, 1942.
45. Rubinshteyn, Ye. S. Trudy Nauchno-Issledovatel'skikh Uchrezhdeniy GUGMS, Series 1, No. 22, 1946.
46. Rubashev, B. M. Byull. KISO, No. 2, 1949.
47. Eygenson, M. S., Gnevyshev, M. N., Ol', A. I., Rubashev, B. M. Solnechnaya aktivnost' i yeye zemnyye proyavleniya (Solar Activity and its Terrestrial Manifestations). Moscow-Leningrad, GTTI, 1948.
48. Eygenson, M. S. Tsirk. LAO, No. 29, 1955.
49. Andreyanov, V. G. Trudy Gosudarstvennogo Gidrologicheskogo Instituta, No. 68, 1959.
50. Predtechenskiy, P. P., Gurevich, B. S. Trudy GGO, Vol. 8, No. 70, 33, 1948.
51. Girs, A. A. Meteorologiya i Gidrologiya, No. 12, 1958.
52. Soskin, I. I. Trudy Okeanograficheskoy Komissii AN SSSR, Vol. 7, 3, 1960.
53. Ringnes, T. S. Nature, Vol. 192, 151, 1961.
54. Bell, B. Publ. Astr. Soc. Pacif., Vol. 71, 165, 1959.
55. Gleissberg, W. Naturwissenschaften, Vol. 42, 410, 1955.
56. Maksimov, I. V. Doklady Akademii Nauk SSSR, Vol. 42, 1149, 1952.
57. Schove, D. J. Journal of Geophysical Research, Vol. 60, 127, 1955.
58. Eygenson, M. S. Dokl. Akad. Nauk SSSR, Vol. 93, 413, 1946.
59. --- Astron. Zh., Vol. 25, 101, 1948.
60. Kopecky, M. Publications of the Astronomical Institute of the Czechoslovakian Academy of Sciences, No. 42, 1958.
61. --- Postepy astron., Vol. 6, 39, 1958.
62. Clough, H. W. Ap. J., Vol. 22, 42, 1905.
63. --- Monthly Weather. Rev., Vol. 60, 99, 1933.
64. Köppen, W. Meteorologische Zeitschrift, Vol. 47, 205, 1930.
65. Cullmer, C. J. Smith. Misc. Coll., Vol. 103, No. 10, 1943.

66. Bezrukova, A. Ya. Byull. KISO, No. 7, 21, 1951.
67. Bezrukova, N. Ya. Solnechnyye Dannyye, No. 2, 1957.
68. Brunner-Hägger, W., Liepert, A. Astr. Mitt., No. 40, 1941.
69. Brunner-Hägger, W. Astr. Mitt., No. 144, 1945.
70. Gleissberg, W. Publ. Istanbul. Univ. Obs., No. 30, 1946.
71. Kermack, W. O., McKendrick, A. G. Proceedings of the Royal Society (Edinburgh), A, Vol. 57, 228, 1937.
72. Xanthakis, J. Comptes Rendus, Vol. 253, 1311, 1961.
73. Anderson, C. N. Journal of Geophysical Research, Vol. 59, 455, 1954.
74. Herrinck, P. Nature, Vol. 184, 51, 64, 655, 1959.
75. Eygenson, M. S. Ocherki fiziko-geograficheskikh proyavleniy solnechnoy aktivnosti (Essays on the Physicogeographic Manifestations of Solar Activity). L'vovskiy universitet, 1957.
76. Bonev, N. Soln. Dannyye, No. 3, 1957.
77. Dzhurkovich, P. Soln. Dannyye, No. 6, 1956.
78. Ol', A. I. Byull. KISO, No. 1, 1949.
79. Shnitnikov, A. V. Byull. KISO, No. 7, 1951.
80. Denning, W. F. Observatory, Vol. 38, 324, 1915.
81. Alter, D. Ann. Math. Statistics, Vol. 8, 121, 1937.
82. Schostakowitsch, W. B. Gerlands Beiträge zur Geophysik, Vol. 30, 281, 1931.
83. Willett, H. C. Climatic Change. Cambridge, Harvard, 1953.
84. Predtechenskiy, P. P. Priroda, No. 6, 1946.
85. Schuster, A. Proc. Roy. Soc. (London), A, Vol. 77, 141, 1906.
86. Labrouste, H., Labrouste, J. Third report of the Commission of Solar-terrestrial relations. Paris, 1931.
87. Vitel's, L. A. Meteorologiya i Gidrologiya, No. 6, 1946.
88. Eygenson, M. S., Prokof'yeva, I. A. Byull. KISO, No. 5-6, 1950.
89. Nikol'skiy, A. P. Problemi Arktiki, No. 4, 1938.
90. Baur, F. Meteorologische Rundschau, Vol. 2, No. 1, 1949.
91. --- Archiv für Meteorologie, Geophysik, und Bioklimatologie, A, Vol. 6, 129, 1953.
92. Rima, A. Geofisica e Meteorologia, Vol. 9, 39, 1961.
93. Link, F., Ceplecha, Z. BAC, Vol. 1, 61, 1949.
94. Unterweger, J. Denkschrift der Wiener Akademie der Wissenschaft, Vol. 58, 55, 1891.
95. Frenkel, E. Publikationen der Sternwarte Zürich, V, 47, 1913.
96. Vitinskiy, Yu. I. Soln. Dannyye, No. 7, 1957.
97. Eygenson, M. S., Rubashev, B. M. Astron. Zh., Vol. 17, 17, 1940.
98. De Jager, C. Vistas in Astronomy, 4, London, Pergamon Press, 1961; Stroyeniye i dinamika atmosfery Solntsa (Structure and Dynamics of the Solar Atmosphere). Moscow, IL, 1962.
99. Lockyer, N., Lockyer, W. Proc. Roy. Soc. (London), Vol. 70, 500, 1902.

100. Drozdov, O. A. Trudy GGO, Vol. 19, No. 81, 102, 1950.
101. Newcomb, S. Ap. J., Vol. 13, 1, 1901.
102. Becker, U. Z. Astrophys., Vol. 48, 88, 1959.
103. Kopecky, M. BAC, Vol. 11, 158, 1960.
104. --- BAC, Vol. 13, 63, 1962.
105. Mascari, A. Mem. Soc. spettr. Ital., Vol. 33, 45, 1904.
106. D'Azambuja, L. Ann. Obs. Paris Meudon, Vol. 6, No. 47, 1948.
107. Pettit, E. Astrophysics, New York, 1951.
108. Carrington, R. C. MN, Vol. 19, 1, 1858.
109. Spörer, G. Publications of the Astronomical Observatory Potsdam, Vol. 10, 144, 1894.
110. Maunder, E. W. MN, Vol.82, 534, 1922.
111. Lockyer, W. S. J. MN, Vol. 64, 5, 1904.
112. Bell, B. Smiths. Contr. Astr., Vol. 5, No. 3, 1960.
113. Maunder, E. W. MN, Vol. 64, 747, 1904.
114. Vitinskiy, Yu. I. Soln. Dannyye, No. 1, 1961.
115. Spörer, G. Publ. Astr. Obs. Potsdam, Vol. 2, 1881.
116. Gleissberg, W. Publ. Istanbul. Univ. Obs., No. 20, 1942.
117. Gnevyshev, M. N., Gnevysheva, R. S. Byull. KISO, Vol. 1, 1949.
118. Rubashev, B. M. Izvestiya GAO, No. 162, 1958.
119. Gleissberg, W. Ap. J., 219, 1944.
120. Rubashev, B. M. Byull. KISO, No. 10, 1953.
121. Carrington, R. C. MN, Vol. 19, 81, 1859.
122. Spörer, G. Publ. Astr. Obs. Potsdam, Vol. 4, 415, 1889.
123. Tuominen, J. Z. Astrophys., Vol. 30, 261, 1952.
124. --- Vistas in Astronomy, Vol. 1, London—N. Y., Pergamon Press, 1955.
125. Schwarzschild, M., Richardson, R. S. Convegnodi Volta, Roma, 1953.
126. Becker, U. Z. Astrophys., Vol. 34, 129, 1954.
127. Tuominen, J. Z. Astrophys., Vol. 51, 91, 1961.
128. Vitinskiy, Yu. I. Soln. Dannyye, No. 3, 1960.
129. Maunder, E. W. MN, Vol. 79, 457, 1919.
130. Bezrukova, A. Ya. Byull. KISO, No. 7, 1951.
131. --- Izv. GAO, No. 159, 1958.
132. Vsekhsvyatskiy, S. K. Byull. KISO, No. 5-6, 1950.
133. Abetti, G. Boll. geod. e sci. affini, Vol. 16, 635, 1957.
134. Krat, V. A. Figury ravnovesiya nebesnykh tel (Figures of Equilibrium of Celestial Bodies). Moscow-Leningrad, GTTI, 1950.
135. Waldmeier, M. Z. Astrophys., Vol. 26, 147, 1949.
136. Bruzek, A. Z. Astrophys., Vol. 51, 75, 1961.
137. Aslanov, I. A. Avtoreferat kandidatskoy dissertatsii (Author's Abstract of Candidate's Thesis). Leningrad, 1962.
138. Cimino, C. Convegno di Volta, Roma, 1953.
139. Jannuzi, M. A. Atti Accad. naz. Lincei Mem. Cl. sci. fis. mat. e natur., ser. I, Vol. 2, 277, 1955.
140. Gething, P. J. D. MN, Vol. 115, 558, 1955.

141. Jannuzzi, M. A. Mem. Soc. Astr. Ital., Vol. 26, 447, 1955.
142. Schur, W., Ambronn, L. Göttingen Abhandlungen, Vol. 3, No. 3, 1905.
143. Meyermann, B. AN, Vol. 269, 114, 1939.
144. Wolfer, A. Publ. Obs. Zürich, Vol. I, 317, 1897; Mem. soc. spett. ital., Vol. 29, 152, 1900.
145. Maunder, E. W. MN, Vol. 65, 2, 1904; Vol. 66, 538, 1905.
146. Bartels, J. Terrestrial Magnetism, Vol. 37, 1, 1932.
147. Rubashev, B. M. Priroda, No. 8, 1947.
148. --- Byull. KISO, No. 2, 1949.
149. Deslandres, H. Comptes Rendus, Vol. 182, 1301, 1926; Vol. 183, 165, 493, 1313, 1926; Vol. 185, 10, 1626, 1927.
150. Salet, P. J. Obs., Vol. 16, 213, 1933.
151. Brooks, C. E. P., Carruthers, N. Statistical Methods in Meteorology, H. M., London, 1953.
152. Becker, U. Z. Astrophys., Vol. 37, 47, 1955.
153. Vitinskiy, Yu. I. Soln. Dannyye, No. 3, 1958.
154. Losh, H. M. Publ. Obs. Michigan, Vol. 7, 127, 1938.
155. Cortie, A. L. MN, Vol. 73, 146, 1913.
156. Predtechenskiy, P. P., Brodovitskiy, K. V. Trudy Tashkentskoy Geofizicheskoy Observatorii, No. 1, 1940.
157. Rubashev, B. M. Kandidatskaya dissertatsiya (Candidate's Dissertation). 1948.
158. Vitinskiy, Yu. I. Izv. GAO, No. 163, 1960.
159. Abetti, G. Handbuch der Astrophysik, IV, Berlin, Springer, 1929.

Chapter 2

1. Kukarkin, B. V. Peremennyye zvezdy, I (Variable Stars, Part I). Moscow-Leningrad, GTTI, 1937.
2. Krank-Kamenetskiy, D. A. Fizicheskiye protsessy vnutri zvezd (Physical Processes Within Stars). Moscow, Fizmatgiz, 1959.
3. Schwarzschild, M. Stroyeniye i evolutsiya zvezd (Structure and Evolution of Stars). Moscow, IL, 1961.
4. Masevich, A. G., Volkonskaya, T. G. Astron. Zh., Vol. 37, 42, 1960.
5. Waldmeier, M. Ergebnisse und Probleme der Sonnenforschung (Results and Problems of Solar Investigations). Leipzig, Geest und Portig, 1955.
6. Oster, L. Z. Astrophys., Vol. 42, 228, 1957.
7. Vasil'yeva, G. Ya. Izv. GAO, No. 170, 1962.
8. Kuiper, G. Solntse (The Sun). Moscow, IL, 1957.
9. Krat, V. A. Figury ravnovesiya nebesnykh tel (Figures of Equilibrium of Celestial Bodies). Moscow-Leningrad, GTTI, 1950.
10. Pikel'ner, S. B. Osnovy kosmicheskoy elektrodinamiki (Principles of Cosmic Electrodynamics). Moscow, Fizmatgiz, 1961.

11. Wasiutynski, I. Astrophys. norveg., Vol. 4, 1, 1946.
12. Kolmogorov, A. N. Dokl. Akad. Nauk SSSR, Vol. 30, 299, 1941.
13. Obukhov, A. M. Dokl. Akad. Nauk SSSR, Vol. 32, 22, 1941.
14. Vasil'yeva, G. Ya. Soln. Dannyye, No. 4, 1963.
15. Yudin, M. I., Gandin, L. S., Laykhtman, D. L., Matveyev, L. G. Dinamicheskaya meteorologiya (Dynamic Meteorology). Leningrad, Gidrometizdat, 1955.
16. Krat, V. A. Byulleten' Observatorii im. Engel'gardta, No. 11, 1937.
17. Csada, I. K. Contrib. Konkoly, No. 19, 1948.
18. Schmeidler, F. Zeitschrift für Naturforschung, Vol. 5a, 297, 1950.
19. Rubashev, B. M. Izv. GAO, No. 162, 1958.
20. Csada, I. K. Soln. Dannyye, No. 6, 1956.
21. Schoenberg, E. Annalen der Bayrischen Akademie der Wissenschaften, Vol. 85, 1, 1958.
22. Broxson, I. W. Physical Review, Vol. 62, 521, 1942.
23. Nicholson, S. Publs. Astron. Soc. Pacit., Vol. 45, 51, 1933.
24. Hautgast, J., Sluiters, A. Bulletin of the Astronomical Institute, The Netherlands, Vol. 108, 354, 1948.
25. Mattig, W. Z. Astrophys., Vol. 31, 273, 1953.
26. Ikhsanov, R. N. Soln. Dannyye, No. 12, 1961.
27. Hale, G. E. MN, Vol. 82, 168, 1922.
28. Cowling, T. Magnitnaya gidrodinamika (Magnetohydrodynamics). Moscow, IL, 1959.
29. Alfvén, H. Kosmicheskaya elektrodinamika (Cosmic Electrodynamics). Moscow, IL, 1952.
30. Elsasser, W. Phys. Rev., Vol. 69, 106, 1946.
31. Bullard, E. Proc. Roy. Soc. (London), A, Vol. 197, 133, 1949.
32. Parker, E. Ap. J., Vol. 122, 293, 1955.
33. Schwarzschild, K. Göttingen Nachrichten, Vol. 41, 1906.
34. Rayleigh, W. Philos. Mag., Vol. 32, 529, 1916.
35. Benard, H. Annales des Chimie et Physique, Vol. 23, 62, 1901.
36. Krat, V. A. Izv. GAO, No. 170, 1962.
37. Unsöld, A. Physik der Sternatmosphären (Physics of Star Atmospheres). Berlin, Springer, 1955.
38. Schmidt, W. Wiener Sitzungsberichte, Vol. 11a, 126, 1917.
39. Rakipova, L. R. Teplovoy rezhim atmosfery (Heat Regime of the Atmosphere). Leningrad, Gidrometizdat, 1957.
40. Schmeidler, F. AN, Vol. 283, 66, 1956.
41. Taylor, G. J. Philosophical Transactions of the Royal Society of London, A, Vol. 223, 289, 1922.
42. Gandin, A. S. Dokl. Akad. Nauk SSSR, Vol. 55, 705, 1947.
43. Chandrassekhar, S. Proc. Roy. Soc., A, Vol. 217, 306, 1953.
44. --- Philos. Mag., Vol. 43, 501, 1952; Vol. 45, 1177, 1954.
45. Vitense, E. Z. Astrophys., Vol. 32, 135, 1953.
46. Unsöld, A. Z. Astrophys., Vol. 2, 209, 1931.
47. Siedentopf, H. AN, Vol. 247, 297, 1933.
48. Waldmeier, M. Einführung in die Astrophysik (Introduction to Astrophysics). Birkhäuser, Basel, 1948.

49. Biermann, L. Z. Astrophys., Vol. 21, 32, 1942.
50. Strömgren, B. Mat. Tidsskr., Vol. 96, 1950.
51. Böhm, K. H. Z. Astrophys., Vol. 46, 245, 1958.
52. Lebedinskiy, A. I. Uchenyye Zapiski LGU, Vol. 11, 17, 1941.
53. Böhm, K. H., Richter, E. Z. Astrophys., Vol. 48, 231, 1959; Vol. 50, 79, 1960.
54. Biermann, L., Kippenhahn, K., Temesvary, St. Z. Astrophys., Vol. 48, 172, 1959.
55. Skumanich, A. Ap. J., Vol. 121, 408, 1955.
56. Epstein, J., Motz, L. Ap. J., Vol. 120, 126, 1954.
57. Biermann, L., Temesvary, A. AN, Vol. 283, 60, 1956.
58. Emden, R. Gaskugeln (Gaseous Spheres). Teubner, Leipzig, 1907.
59. Weyman, R. Ap. J., Vol. 126, 208, 1957.
60. Searce, R. Ap. J., Vol. 129, 489, 1959.
61. Schwarzchild, M., Howard, R., Härm, R. Ap. J., Vol. 125, 233, 1957.
62. Rosseland, S. Astrophys. norveg., Vol. 2, 173, 1936.
63. Jeans, J. Astronomy and Cosmogony. Cambridge University Press, 1929.
64. Wasiutynski, I. Ann. astrophys., Vol. 21, 119, 1958.
65. Abetti, G. Handbuch der Astrophysik, IV Springer, Berlin, 1929.
66. Bjerknes, V. Astrophys. norveg., Vol. 2, 263, 1937.
67. Rosseland, S. Ap. J., Vol. 63, 364, 1926.
68. Krogdahl, W. Ap. J., Vol. 99, 191, 1944.
69. Randers, G. Astrophys. norveg., Vol. 3, 97, 1939.
70. Klauder, H. AN, Vol. 250, 89, 1933.
71. Krat, V. A. Izv. GAO, No. 136, 1947.
72. Izvekov, B. I. Dinamicheskaya meteorologiya (Dynamic Meteorology). Vols. 1 and 2, Leningrad, Gidrometizdat, 1937.
73. Brent, D. Fizicheskaya i dinamicheskaya meteorologiya (Physical and Dynamic Meteorology). Leningrad, Gidrometizdat, 1938.
74. Eddington, A. S. MN, Vol. 30, 54, 1929.
75. Vogt, H. AN, Vol. 223, 229, 1925.
76. Milne, E. Handbuch der Astrophysik, III. Springer, Berlin, 1930.
77. Sweet, P. A. MN, Vol. 110, 69, 1950.
78. Csada, I. K. Izvestiya Astronomicheskogo instituta konkol'ya, No. 22, 1949.
79. Kochin, N. Ye. Sobraniye Sochineniy (Collection of Papers). Vol. I, 286, 1949.
80. Rubashev, B. M. Byull. KISO, No. 10, 1953.
81. Krat, V. A. Uspekhi Astronomicheskikh Nauk, Vol. 3, 129, 1947.
82. Rubashev, B. M. Byull. KISO, No. 8-9, 1953.
83. Gnevysheva, R. S. Astron. Zh., Vol. 18, 27, 1941.
84. Beckers, J. M. Bull. Astron. Inst. Netherl., Vol. 15, 85, 1960.
85. Tuominen, J. Observatory, Vol. 66, 160, 1945.
86. Plaskett, H. H. MN, Vol. 119, 197, 1959.
87. Prandtl, L. The Essentials of Fluid Dynamics. London-Glasgow, Blackie, 1953.

88. Kippenhahn, R. Z. Astrophys., Vol. 38, 166, 1955.
89. --- Z. Astrophys., Vol. 46, 26, 1958.
90. Baker, N., Kipenhahn, R. Z. Astrophys., Vol. 48, 140, 1959.
91. Kippenhahn, R. Z. Astrophys., Vol. 48, 203, 1959.
92. Hoyle, F., Wickramsinghe, N. C. MN, Vol. 123, 51, 1961.
93. Rubashev, B. M. Soln. Dannyye, No. 10, 1959.
94. Biermann, L. Symposium IAU, No. 6, 1956.
95. Gleissberg, W. Ap. J., Vol. 100, 219, 1944.
96. Alfvén, H. Tellus, Vol. 8, 274, 1956.

Chapter 3

1. Biermann, L., Lüst, R. Mem. Soc. Roy. Sci. Liège, Vol. 4, 623, 1961.
2. Gnevyshev, M. N., Gnevysheva, R. S. Byull. KISO, No. 1, 1949.
3. Waldmeier, M. Astr. Mitt., No. 138, 1939.
4. Rubashev, B. M. Izv. GAO, No. 162, 1958.
5. Gleissberg, W. Ap. J., Vol. 100, 219, 1944.
6. Biermann, L. Z. Naturforsch., Vol. 7a, 103, 1952.
7. Alfvén, H. Tellus, Vol. 8, 274, 1956.
8. Rubashev, B. M. Soln. Dannyye, No. 11, 1960.
9. Ikhsanov, R. N. Soln. Dannyye, No. 12, 1961.
10. Kuiper, G. Solntse (The Sun). Moscow, IL, 1957.
11. Walen, C. Arkiv mat. astron., fys., Vol. 33A, No. 18, 1946.
12. Pikel'ner, S. B. Osnovy kosmicheskoy elektrodinamiki (Principles of Cosmic Electrodynamics). Moscow, Fizmatgiz, 1961.
13. Walen, C. Arkiv mat. astron., fys., Vol. 30A, No. 15, 1944.
14. Anderson, C. N. Journal of Geophysical Research, Vol. 59, No. 4, 1954.
15. Link, F. BAC, Vol. 1, 61, 1949.

Chapter 4

1. Chapman, S. Kosmicheskaya astrofizika (Cosmic Astrophysics). Moscow, IL, 1962.
2. Kolokolov, V. P. Trudy GGO, Vol. 35, 21, 1952.
3. Megrelishvili, T. G. Izv. AN SSSR, Seriya Geofizicheskaya, No. 8, 1956; No. 6, 1959.
4. Rubashev, B. M. Soln. Dannyye, No. 6, 1956.
5. Mitra, S. K. Verkhnyaya atmosfera (The Upper Atmosphere). Moscow, IL, 1955.
6. Rubashev, B. M. Soln. Dannyye, No. 11, 1961.
7. Gerson, N. S. Uspekhi Fizicheskikh Nauk, Vol. 47, 616, 1952.
8. Stranz, D. Tellus, Vol. 2, 150, 1950.
9. Ivanov-Kholodnyy, G. S. Izv. AN SSSR, Ser. Geofiz., No. 1, 1959.

10. Chosh, S. N., Nand, S. Indian Journal of Physics, Vol. 34, 516, 1960.
11. Chapman, S. Terr. Magn., Vol. 42, 355, 1937.
12. --- MN, Vol. 79, 70, 1919.
13. Shklovskiy, I. S. Izvestiya KrAO, Vol. 8, 51, 1952.
14. Paetzold, H. K. J. Geophys. Res., Vol. 67, 2741, 1962.
15. Omholt, A. Journal of Atmospheric and Terrestrial Physics, Vol. 5, 243, 1954.
16. Nicolet, M. Annales Geophysique, Vol. 15, 1, 1959.
17. Kamiyama, T. Tohoku Univ. Sci. Rep., ser. 5, Vol. 9, 2, 1957.
18. Rakipova, L. R. Byull. KISO, No. 7, 1951.
19. Nikol'skiy, A. P. Soln. Dannyye, No. 6, 1955.
20. Lettau, H. Atmosphärische Turbulenz (Atmospheric Turbulence). Leipzig Akademie Verlag, 1939.
21. Johnson, F. S. J. Geophys. Res., Vol. 65, 2227, 1960.
22. Cole, K. D. Nature, Vol. 194, 75, 1962.
23. --- Australian Journal of Physics, Vol. 15, 223, 1962.
24. King-Hele, D. G. Nature, Vol. 194, 39, 1962.
25. Kallman-Bijl, H. K. J. Geophys. Res., Vol. 66, 787, 1961.
26. Kondrat'yev, K. Ya., Filippovich, O. P. Teplovoy rezhim verkhnikh sloyev atmosfery (Heat Regime of the Upper Layers of the Atmosphere). Leningrad, Gidrometizdat, 1960.
27. Paetzold, H. K. J. Geophys. Res., Vol. 67, 2741, 1962.

Chapter 5

1. Eygenson, M. S., Gnevyshev, M. N., Ol', A. I., Rubashev, B. M. Solnechnaya aktivnost' i yeye zemnyye proyavleniya (Solar Activity and its Terrestrial Manifestations). Moscow-Leningrad, GTTI, 1948.
2. Rubashev, B. M. Trudy Vsesoyuznogo Nauchno-Meteorologicheskogo Soveshchaniya 1961 (Transactions of the All-Union Scientific Meteorological Conference of 1961). Vol. IV, 252, 1962.
3. Köppen, W. Osterreiche Zeitschrift der Meteorologie, Vol. 8, 237, 1873.
4. Wolf, R. Handbuch der Astronomie (Handbook of Astronomy). Vol. 2, Zürich, Schulthess, 1892.
5. Wolf, Chr. Elementare Astronomie (Elementary Astronomy). Leipzig, 1716.
6. Herschel, W. Philos. Trans. Roy. Soc., London, Vol. 101, 1, 1801.
7. Arago, F. Notices Scientifiques. Paris, Barral, 1820.
8. Gautier, A. Ann. Chimie et Phys., Vol. 12, 57, 1844.
9. Wolf, R. Astr. Mitt., Zurich, No. 9, 1859.
10. Nansen, F., Helland-Hansen, B. Videnskapsselskapets Skr., I, Math. naturw. Klasse, No. 9, 1916.
11. Clayton, H. H. Bulletin of the American Meteorological Society, Vol. 14, 65, 1933.

12. Walker, G. Indian Met. Mem. Vol. 21, 11, 1915.
13. Brooks, C. E. P. Quarterly Journal, Royal Meteorological Society, Vol. 45, 233, 1918.
14. --- Quarterly Journal, Royal Meteorological Society, Vol. 60, 153, 1934.
15. Fedorov, Ye. Ye. Izvestiya Glavnoy Fizicheskoy Observatorii, Vol. 1, No. 3, 1920.
16. Vize, V. Yu. Zhurnal Geofiziki i Meteorologii, Vol. 1, 78, 1924.
17. Abbot, C. G., Fowle, F. E. Annals of the Astrophysical Observatory of the Smithsonian Institutions, Vol. 2, 1, 1908.
18. Humphreys, W. Ap. J., Vol. 32, 87, 1910.
19. Memery, H. Bulletin Observatoire, Talence, Vol. 43, 236, 1940.
20. Aufsess, V. Met. Z., Vol. 44, 227, 1927.
21. Brodovitskiy, K. V., Predtechenskiy, P. P. Trudy Tashkentskoy Geofizicheskoy Observatorii, Vol. 1, 1, 1940.
22. Tikhomirova, M. F. Zhurnal Geofiziki i Meteorologii, Vol. 6, 153, 1929.
23. Rubashev, B. M. Doklady AN SSSR, Vol. 27, No. 8, 792, 1940.
24. Veynberg, B. P. Geofizicheskiy Sbornik, Vol. 7, No. 1, 1, 1929.
25. Gleissberg, W. Publ. Istanbul Univ. Obs., No. 30, 1946.
26. Rubashev, B. M. Byull. KISO, No. 2, 1949.
27. Ol', A. I. Soln. Dannyye, No. 1, 1954.
28. Krat, V. A. Izv. GAO, No. 139, 1, 1947.
29. Cirera, R., Balcells, M. Comptes Rendus, Vol. 144, 159, 1907.
30. Prokof'yeva, I. A. Byull. KISO, No. 1, 1949.
31. --- Atmosfernyy ozon (Atmospheric Ozone). Moscow-Leningrad, Izd-vo AN SSSR, 1951.
32. Dobson, G. M. B., Harrison, D. N. Proc. Roy. Soc., A, 129, 411, 1930.
33. Willet, H. J. Geophys. Res., Vol. 67, 661, 1962.
34. Mitchell, J. M. J. Geophys. Res., Vol. 67, 4093, 1962.
35. Fritz, S. Proceedings of the National Academy of Science, USA, Vol. 43, No. 1, 1957.
36. Haurwitz, B. Cent. Symp. Harward, No. 7, 1948.
37. Vassy, A., Vassy, E. Reunion de la comission mixte pour l'etude des relations entre les phenomenes solaires et terrestres (Meeting of the Mixed Commission for the Study of the Relationship between Solar and Terrestrial Phenomena). Paris, September 3, 1952.
38. Eygenson, M. S. Ocherki fiziko-geograficheskikh proyavleniy solnechnoy aktivnost' (Essays on Physicogeographic Manifestations of Solar Activity). Izd-vo L'vovskogo universiteta, 1957.
39. Gnevyshev, M. N. Izv. GAO, No. 131, 1938.
40. Kuiper, G. Solntse (The Sun). Moscow, IL, 1957.
41. Hanzlick, S. Gerlands, Beiträge zur Geophysik, Vol. 29, 138, 1931.
42. Easton, V. Met. Z., Vol. 45, 452, 1928.
43. Köppen, W. Met. Z., Vol. 47, 205, 1930.
44. Scherhag, R. Annalen der Hydrographie, Vol. 64, 397, 1936.

45. Scherhag, R. Neue Methoden der Wetteranalyse und Wetterprognose (New Methods of Weather Analysis and Forecasts). Springer, Berlin, 1948.
46. Gleissberg, W. Häufigkeit der Sonnenflecken (Frequency of Sunspots). Akademie Verlag, Berlin, 1952.
47. Angström, A. Geogr. ann., Vol. 21, 119, 1939.
48. Defant, A. Geogr. ann., Vol. 3, 200, 1921.
49. Rakipova, L. R. Teplovoy rezhim atmosfery (Heat Regime of the Atmosphere). Leningrad, Gidrometizdat, 1957.
50. Duell, B., Duell, H. Smith. Misc. Coll., Vol. 110, No. 8, 1948.
51. Koppe, H. Annalen der Meteorologie, Vol. 4, 87, 1951.
52. Willett, H. C. Annals of the New York Academy of Science, Vol. 95, 89, 1961.
53. Craig, R. A. J. Meteorol., Vol. 9, 126, 1952.
54. Flohn, H. Arch. Meteorol. Geophys. und Bioklimatol., Vol. A 3, 303, 1951.
55. Septer, E. Met. Z., Vol. 43, 229, 1926.
56. Vitel's, L. A. Byull. KISO, No. 1, 1949.
57. Valnicek, B. BAC, Vol. 3, 71, 1952; Vol. 4, 97, 1953.
58. Krjivsky, BAC, Vol. 2, 161, 1951.
59. Prudhomme, A. Meteorologie, Vol. 23, 185, 1951.
60. Brier, G. Technical Reports, Institute of Solar-Terrestrial Research, No. 1, 1956.
61. Woodbridge, D. D., Porte, T. W., Macdonald, N. G. Techn. Reptr., Inst. Solar.-Terr. Res., No. 3, 1957.
62. Macdonald, N. G., Roberts, W. O. J. Geophys. Res., Vol. 65, 529, 1960.
63. Rodevald, M. Ann. Meteorol., Vol. 7, 186, 1955.
64. Shapiro, R. J. Meteorol., Vol. 13, 335, 1956.
65. London, J., Ruff, J., Tickl, L. J. Geophys. Res., Vol. 64, 1827, 1959.
66. Ward, F. Techn. Reports Inst. Solar-Terrestrial. Res., No. 1, 1956.
67. Wurlitzer, G. Meteorologische Abhandlungen des Instituts für Meteorologie und Geophysik, Freie Universität, Berlin, Vol. 7, 47, 1958.
68. Graics, A. Idöjares, Vol. 63, 236, 1959.
69. Gabilly, A. Meteorologie, Vol. 50, 109, 1958.
70. Rubashev, B. M. Soln. Dannyye, No. 5, 1957.
71. Romanovskiy, V. I. Primeneniye matematicheskoy statistiki v opytnom dele (Application of Mathematical Statistics in Experimentation). Moscow, GITTL, 1947.
72. Drozdov, O. A., Pokrovskaya, T. V. Trudy Vsesoyuznogo Nauchno-Meteorologicheskogo Soveshchaniya 1961, Vol. IV, Leningrad, Gidrometizdat, 1962.
73. Rubashev, B. M. Soln. Dannyye, No. 6, 1957.
74. Abbot, C. G. Smith. Misc. Coll., Vol. 107, No. 3, 1947.

75. Abbot, C. G. Smith. Misc. Coll., Vol. 110, No. 4, 1948.
76. --- Smith. Misc. Coll., Vol. 111, No. 5, 1949.
77. Vitel's, L. A. Trudy Tsentral'nogo Instituta Prognozov, Vol. 51, 22, 1957.
78. --- Trudy GGO, Vol. 87, 56, 1959.
79. --- Soln. Dannyye, No. 4, 1955.
80. Visser, S. W. Transactions of the American Geophysical Union, Vol. 39, 835, 1958.
81. Rangardjean, C. Indian J. Meteorol. Geophys., Vol. 5, 291, 1954.
82. Vitel's, L. A. Trudy GGO, Vol. 8, 51, 1948.
83. Bezrukova, A. Ya. Byull. KISO, No. 5-6, 1950.
84. --- Trudy Laboratorii Ozerovedeniya, Vol. 3, 23, 1954.
85. Vitel's, L. A. Trudy GGO, Vol. 90, 95, 1960.
86. Girs, A. A. Meteorologiya i Gidrologiya, No. 8, 1956.
87. --- Meteorologiya i Gidrologiya, No. 8, 1960
88. Kolobov, N. V. Uchenyye Zapiski Kazanskogo Universiteta, Vol. 117, 333, 1957.
89. Dzerdzeyevskiy, B. L. Trudy Leningradskogo Gidrometeorologicheskogo Instituta, No. 5-6, 191, 1956.
90. Mishkarev, A. I. Meteorologiya i Gidrologiya, No. 12, 26, 1959.
91. Ped', D. A., Sidochenko, T. V. Meteorologiya i Gidrologiya, No. 9, 20, 1959.
92. Horn, W. Küste, Vol. 8, 83, 1960.
93. Soskin, I. M. Trudy Okeanograf. Komissii AN SSSR, Vol. 7, 3, 1960.
94. --- Trudy GGO, Vol. 7, 3, 1960; Vol. 89, 73, 1959.
95. --- Materialy konferentsii "Vzaimodeystviye atmosfery i gidrosfery v severnoy chasti Atlanticheskogo okeana" (Materials of the Conference "Interaction Between the Atmosphere and Hydrosphere in the Northern Part of the Atlantic Ocean"). Vol. 2, 17, 1959.
96. Butorin, N. V. Vekovyye izmeneniya srednego urovnya Atlanticheskogo okeana i ikh svyaz' s tsirkulyatsiyey atmosfery (Secular Changes of the Mean Level of the Atlantic Ocean and Their Relationship to Atmospheric Circulation). Moscow-Leningrad, Izd-vo AN SSSR, 1960.
97. Krapfenbauer, A. Forstzeitung, Vol. 71, 125, 1960.
98. Maksimov, I. V. Dokl. AN SSSR, Vol. 86, 917, 1952.
99. --- Dokl. AN SSSR, Vol. 92, 1149, 1953.
100. Richter-Bernburg, A. Naturwissenschaften, Vol. 37, 11, 1950.
101. Anderson, C. N. J. Geophys. Res., Vol. 59, 455, 1954.
102. Link, F., Linkova, Z. Studia geophys. et geod., Vol. 3, 43, 1959.
103. Rodewald, M. Verhandl. Deutsche Geographentag, Vol. 31, 215, 1958.
104. Flohn, H. Verhandl. Deutsche Geographentag, Vol. 31, 201, 1958.
105. Fairbridge, R. W. Scientific American, Vol. 202, 70, 1960.
106. Schove, D. J. J. Geophys. Res., Vol. 60, 127, 1955.
107. Shnitnikov, A. V. Publ. Assoc. Internat. Hydrol. scientist, No. 51, 1960.

108. Willett, H. C. J. Meteorol., Vol. 6, 34, 1949.
109. Predtechenskiy, P. P. Byull. KISO, No. 5-6, 1960.
110. Shvartsbakh, M. Klimaty proshlogo (Climates of the Past). Moscow, IL, 1955.
111. Predtechenskiy, P. P. Trudy GGO, Vol. 19, 193, 1950.
112. Brooks, C. E. P. Klimaty proshlogo (Climates of the Past). Moscow, IL, 1952.
113. Milankovich, M. Matematicheskaya klimatologii i astronomicheskaya teoriya kolebaniy klimata (Mathematical Climatology and the Astronomical Theory of Climatic Variations). Moscow-Leningrad, GTTI, 1939.
114. Flohn, H. Geol. Rundschau, Vol. 40, No. 1, 1952.
115. Viete, G. Tellus, Vol. 2, 102, 1950.
116. Wexler, H. Climatic Change. Cambridge, Harvard, 1953.
117. --- Trans. N. Y. Acad. Sci., Vol. 13, 282, 1951.
118. Goldschmidt, H., Reuter, A. Z. Meteorol., Vol. 3, No. 1/2, 1949.
119. Carapiperis, L. N. Gerlands Beitr. Geophys., Vol. 63, No. 2, 149, 1951.
120. Godoli, G. Annali di Geofis., Vol. 6, 125, 1953.
121. Polli, S. Geof. pura e appl., Vol. 25, 130, 1953.
122. Myrbach, O. Met. Z., Vol. 52, 225, 1935.
123. Moiseyev, A. P. Priroda, No. 1, 1948.
124. Vitel's, L. A. Byull. KISO, No. 7, 1951.
125. Gregor, Z., Krzhivskiy, L. Trudy Chekhoslovatskoy Astrofizicheskoy Observatorii v Ondrzheyove, No. 23, 1953.
126. Bodurtha, F. T., Jr. J. Meteorol., Vol. 9, 118, 1952.
127. Belinskiy, N. A., Glagoleva, M. G. Meteorologiya i Gidrologiya, No. 1, 1959.
128. Brier, G. Techn. Reports Inst. Solar-Terrestr. Res., No. 1, 1956.
129. Dammann, W. Naturwissenschaften, Vol. 47, 529, 1960.
130. --- Arch. Meteorol., Geophys. und Bioklimatol., B, Vol. 10, 17, 1959.
131. Sen-Gupta, P. K. Weather, Vol. 15, 130, 1960.
132. Stranz, D. J. Atmosph. and Terr. Phys., Vol. 16, 180, 1959.
133. Peisino, G. Geofis. e meteorol., Vol. 4, 8, 1956.
134. Maksimov, I. V. Problemy Arktiki, No. 2, 5, 1956.
135. Shnitnikov, A. V. Izvestiya Vsesoyuznogo Geograficheskogo Obshchestva, No. 1, 1960.
136. Willett, H. C. Cent. Proc. Roy. Met. Soc., Vol. 1, 195, 1950.
137. --- J. Meteorol., Vol. 8, 1, 1951.
138. Krjivsky, L. Publications of the Astrophysical Observatory, Ondřjeov, No. 23, 1953.
139. ---BAC, Vol. 1, 147, 1949.
140. Simojoki, H. Geofis. e meteorol., Vol. 6, 25, 1957.
141. Xanthakis, J. Arch. Meteorol. Geophys. und Bioklimatol., A, Vol. 9, 54, 1956.

142. Brückner, E. Klimaschwankungen Seit 1700 (Climatic Variations Since 1700). Geographische Abhandlungen, IV, No. 2, Wien, 1890.
143. Eygenson, M. S. Byull. KISO, No. 7, 1951.
144. Wagner, A. Klimaänderungen und Klimachwankungen (Climatic Changes and Variations). Braunschweig, 1940.
145. Shnitnikov, A. V. Byull. KISO, No. 3-4, 1949; No. 7, 1951.
146. Cullmer, C. J. Smith. Misc. Coll., Vol. 103, No. 10, 1943.
147. Lunsgersgauzen, G. F. Dokl. AN SSSR, Vol. 108, 707, 1956.
148. Newton, H. W., Milsom, A. S. J. Geophys. Res., Vol. 59, 203, 1954.
149. Baur, F. Naturwissenschaften, Vol. 48, 61, 1961.
150. --- Die Sommerniederschläge Mitteleuropas in den letzen 1-1/2 Jahrhunderten und ihre Beziehungen zum Sonnenfleckenzyklus (Summer Rains in Middle Europe in the Last One and one-half Centuries and Their Relation to the Sunspot Cycle). Leipzig Akademie Verlag, 1959.
151. --- Arch. Meteorol. Geophys. und Bioklimatol., A, Vol. 6, 129, 1953.
152. --- Einführung in die Grosswetterkunde (Introduction to Long Range Weather Forecasting). Wiesbaden, Dietrich, 1948.
153. --- Z. Astrophys., Vol. 4, 180, 1932.
154. Pokrovskaya, T. V. Trudy GGO, Vol. 65, 8, 1956.
155. Visser, S. W. Geofis. pura e appl., Vol. 43, 302, 1959.
156. Berkesz, Z. Idöjaras, Vol. 63, 193, 1959.
157. Rubashev, B. M. Dokl. AN SSSR, Vol. 26, 780, 1940.
158. --- Soln. Dannyye, No. 11, 1959.
159. --- Soln. Dannyye, No. 6, 1960.
160. Dmitriyev, A. A. Meteorologiya i Gidrologiya, No. 5, 1949.
161. --- Soln. Dannyye, No. 1, 1957.
162. Rakipova, L. R. Trudy GGO, Vol. 28, 5, 1951.
163. --- Tsirkulyar L'vovskoy Astronomicheskoy Observatorii, No. 55, 1955.
164. Kondrat'yev, K. Ya., Fillipovich, O. P. Teplovoy rezhim verkhnikh sloyev atmosfery (Heat Regime of the Upper Layers of the Atmosphere). Leningrad, Gidrometizdat, 1960.
165. Spar, J. Techn. Rept. Inst. Solar-Terrestr. Res., No. 5, 1958.
166. Scherhag, R. Berliner Deutscher Wetterdienst, No. 38, 51, 1952.
167. Palmer, C. E. J. Meteorol., Vol. 10, 302, 1953.
168. Wexler, H. Atmospheric Explorations, Vol. 79, 20, 1958.
169. Peterson, J., Jones, L. N., Schaefer, F. J., Schulte, H. J. Geophys. Res., Vol. 64, 2331, 1959.
170. Rossi, V. Geophysica, Vol. 4, 147, 1953.
171. Fortak, H. Gerlands Beitr. Geophys., Vol. 67, 66, 1958.
172. Vassy, E. Ciel et terre, Vol. 74, 421, 1958.
173. Ney, E. Nature, Vol. 183, 451, 1959.

Chapter 6

1. Bobrovnikoff, N. T. Contr. Perk. Obs., 16, 1942.
2. Dobrovol'skiy, O. V. Byull. KISO, No. 1, 1949.
3. --- Byull. KISO, No. 3-4, 1949.
4. --- Byull. KISO, No. 5-6, 1950.
5. --- Trudy Instituta Astrofiziki AN Tadzhikskoy SSR, VIII, 1, 1961.
6. Houziau, L. L. Battiau Bull. Cl. Sci. Acad. Roy. Belg., 5 ser. 43, Vol. 171, 1957.
7. Link, F., Vanyssek, V. CR, Vol. 225, 1284, 1947.
8. Richter, N. AN, Vol. 281, 241, 1954.
9. Link, F. La Physique des comets (The Physics of Comets). Louvain, 1953.
10. Sekanina, Z. BAC, Vol. 11, 94, 1960.
11. Biermann, L. Z. Astrophys., Vol. 29, 274, 1951.
12. Vogel, H. C. AN, Vol. 108, 23, 1884.
13. Bobrovnikoff, N. T. Lick Observatory Publications, Vol. 17, 309, 1931.
14. Shul'man, B. S. Izv. GAO, No. 140, 1947.
15. Kupo, I. D., Teyfel', V. G. Astronomicheskiy Tsirkulyar, No. 188, 1958.
16. Kozyrev, N. A. Izv. KrAO, Vol. 16, 148, 1956.
17. Link, F. BAC, Vol. 14, N 1, 1962.
18. Korol'kov, D. V., Pariyskiy, Yu. N., Timofeyeva, S. E., Khaykin, S. E. Dokl. AN SSSR, Vol. 149, 65, 1963.
19. Kellogg, W., Sagan, C. Atmosfery Marsa i Venery (Atmospheres of Mars and Venus). Moscow, IL, 1962.
20. Strong, J., Ross, M. D., Moor, C. B. J. Geophys, Res., Vol. 65, 2526, 1960.
21. Lowell, P. Mars i zhizn' na nem (Mars and Life on It). Odessa, Matezis, 1911.
22. Urey, H. C., Brewer, A. W. Proc. Roy. Soc., A, Vol. 241, 37, 1957.
23. Dolfus, A. Annales d'Astrophysique Supplement, No. 4, Vol. 1, 1957.
24. Vaucouleurs, G. Planet. Space Sci., Vol. 2, 26, 1959.
25. Sharonov, V. V. Priroda planet (Nature of the Planets). Moscow, Fizmatgiz, 1958.
26. Schoenberg, E. AN, Vol. 273, 113, 1942.
27. Kuiper, G. Planets and Satellites. Chicago Univ. Press., 1961.
28. Seidel, L. Untersuchungen über die Lichtstarke der Planeten: Venus, Mars, Jupiter und Saturn (Investigation of Light Emission of Planets: Venus, Mars, Jupiter and Saturn). Bay. Akad. München, 1859.
29. Johnson, H. L., Iriarte, B. Bulletin of the Lowell Observatory, Vol. 4, 99, 1959.
30. Müller, G. Publ. Astrophys. Obs. Potsdam, Vol. 8, 197, 1893.

31. Becker, W. Sitz. d. Preuss. Akad. d. Wiss. Phys. Mat. Kl., Vol. 28, 839, 1933.
32. Schoenberg, E. Sitz. d. Preuss. Akad. d. Wiss. Phys. Mat. Kl., Vol. 26, 383, 1931.
33. Kozyrev, N. A. Izv. KrAO, Vol. 12, 169, 1954.
34. Newkirk, G. Planet. Space Sci., Vol. 1, 32, 1959.
35. Denning, W. F. MN, Vol. 59, 76, 1899.
36. St. Williams, A. Observatory, Vol. 23, 176, 1900.
37. Wonaszek, A. Budapest Publ. d. Stern. Kis-Kartal, Vol. 1, 1, 1901.
38. Bakharev, A. M. Byull. VAGO, No. 4, 11, 1948.
39. Hess, S. L., Panofsky, H. P. Bull. Amer. Meteorol. Soc., Vol. 29, 426, 1948.
40. Shayn, G. A. Izvestiya Rossiyskogo Astronomicheskogo Obshchestva, Vol. 18, 243, 1912.
41. Antoniadi, E. M. MN, Vol. 76, 643, 1916.
42. Hess, S. L. J. Meteorol., Vol. 7, 1, 1950.
43. Abbot, C. G., Aldrich, L. B., Hoover, W. H. Smith. Contr. Astrophys., Vol. 3, 23, 1958.
44. Sterne, T. E., Dieter, N. Smith. Contr. Astrophys., Vol. 3, 9, 1958.
45. Barabashev, N. P. Issledovaniye fizicheskikh usloviy na Lune i planetakh (Investigation of Physical Conditions on the Moon and Planets). Izd-vo Khar'kovskogo universiteta, 1952.

Appendix

1. Eygenson, M. S. Tsirkulyar GAO, No. 30, 1940.
2. Eygenson, M. S., Prokof'yeva, I. A. Byulleten' KISO, No. 5-6, 1950.
3. Mandrykina, T. L. L'vovskiy Astronomicheskiy Sbornik, No. 2, 1954.
4. Kopecky, M. Publications of the Astronomical Institute of the Czechoslovakian Academy of Sciences, No. 28, 1958.
5. Bartels, J. Abh. Preuss. Ak. d. Wissenschaften Phys., Math. Kl., No. 12, 1941.

Translated for the National Aeronautics and Space Administration by John F. Holman and Co. Inc.